UNDERSTANDING THE PERCEPTUAL SEGMENTATION OF SITUATIONS VIA EVENT SEGMENTATION THEORY

A Dissertation
Presented to
The Academic Faculty

by

Joel M. Mumma

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Psychology

Georgia Institute of Technology
August 2019

UNDERSTANDING THE PERCEPTUAL SEGMENTATION OF SITUATIONS VIA EVENT SEGMENTATION THEORY

Approved by:

Dr. Francis T. Durso, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Rick Thomas
School of Psychology
Georgia Institute of Technology

Dr. Jamie Gorman
School of Psychology
Georgia Institute of Technology

Dr. Howard Weiss
School of Psychology
Georgia Institute of Technology

Dr. Joseph Magliano
College of Education and Human Development
Georgia State University

Date Approved: July 15, 2019

ACKNOWLEDGEMENTS

I would like to thank my wife, Kara, for her unending support and patience throughout my time in graduate school. I would also like to thank my advisor, Dr. Francis T. Durso, who not only gave me the tools I needed to succeed, but also gave me every opportunity to use them. Lastly, I would like to express my gratitude to my dissertation committee members (Dr. Rick Thomas, Dr. Jamie Gorman, Dr. Howard Weiss, and Dr. Joseph Magliano) for helping me create a product far better than what I could have created alone and to my undergraduate research assistants (Sarah Mosley, Elizabeth Harvey, Jalen Brown, Erica Callahan, Elena Kusiak, and Katharine Harris), who worked tirelessly and dedicatedly on this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ANOVA	Analysis of Variance
β	Standardized regression coefficient
CAPTION	Complexity, Adversity, Positive Valence, Typicality, Importance, Humor, Negative Valence
CI	Confidence Interval
d	Cohen's *d*
DIAMONDS	Duty, Intellect, Mating, Positivity, Negativity, Deception, Sociality
EST	Event Segmentation Theory
f	Cohen's *f*
F	F statistic
ICC	Intraclass Correlation Coefficient
Log_{10}	Logarithm base 10
Log *b*	Response bias
Log *d*	Discriminability
M	Mean
p	p value
$\eta^2_{partial}$	Partial eta squared
r	Pearson's correlation coefficient
SA	Situation Awareness
SD	Standard Deviation
SE	Standard Error

t	t statistic
X^2	Chi-squared statistic
z	z statistic

SUMMARY

Parsing the daily stream of activity into situations is essential for adaptive functioning in everyday life. Central to this imperative is the question of how people form and modify their mental representation of a situation. Across several literatures (i.e., social psychology, engineering psychology, and narrative comprehension), we identified a number of points of agreement about the properties of our mental representations of situations. We argue that Event Segmentation Theory (EST) provides a framework for understanding how these properties coalesce to give rise to our representations of situations.

The goal of the present studies was to understand how the cognitive mechanisms of EST might account for one property in particular - the hierarchical structure of our representations of situations. According to EST, people maintain a hierarchy of "event models" of ongoing activity in working memory, which represent events unfolding simultaneously on different timescales. Event models continually try to predict the near future and are updated in response to prediction error. Updating an event model gives rise to our perception of a "boundary" between events and is what people report during event segmentation tasks. EST posits that the hierarchy of event models in working memory arises from the differential predictive accuracies of coarse-event models (e.g., of situations) and fine-event models (e.g., of shorter events occurring within situations).

We tested this hypothesis by orienting participants to their event models of the situations or of the fine events in a narrative film, either by having them indicate each time a new situation or a new fine event began. Throughout the film, we also assessed

their confidence and predictive accuracy at moments when both variables should depend on the event model being interrogated. Across two studies, we obtained novel support for the general mechanisms of EST but converging evidence that participants only maintained fine-event models of activity, even though we found that their segmentation of the film depended on their orientation. We propose that the fine-grained segmentation of activity may reflect the updating of fine-event models whereas coarser-grained segmentation may instead reflect how people group fine events online, rather than the updating of coarse-event models (e.g., of situations) per se.

CHAPTER 1. INTRODUCTION

Our daily lives are a continuously unfolding stream of activity. We make sense of this stream by parsing it into discrete, meaningful units (Kurby & Zacks, 2008; Zacks, Speer, Swallow, Braver, & Reynolds, 2007; Zacks & Tversky, 2001). Parsing the activity stream into different situations is essential for adaptive functioning (Rauthmann & Sherman, 2016). For example, if a driver is caught unexpectedly in a downpour, the driver ought to recognize that he or she should not continue driving in the same manner as before the downpour began. Similarly, if a stranger, who stopped you to ask for directions, suddenly produces a knife, one ought to recognize that continuing to give the stranger directions is no longer the most appropriate behavior in the situation. In both examples, it is important that one perceives that a new situation has begun; otherwise, continuing one's behavior from the previous situation would be maladaptive. The consequences of failing to recognize the advent of a new situation are as varied as the situations in which we may find ourselves – from those in fictional worlds of narrative text and film, to the mundane (e.g., a surprise birthday party), to those where life itself hangs in the balance (e.g., flying an airplane through a lightning storm).

The present studies attempt to understand how we divide ongoing activity into different situations through the framework of Event Segmentation Theory (EST; Zacks et al., 2007). EST offers a mechanistic account of how and why people segment continuous activity into discrete events, from those lasting seconds (e.g., eating a strawberry or pouring coffee into a cup) to those lasting tens of minutes (e.g., washing dishes or talking on the phone with a friend; Kurby & Zacks, 2008; Zacks et al., 2007). EST explains the segmentation of activity in terms of the mental representations people have of events (i.e., “event models”; Radvansky, 2012) and the cognitive processes that form and modify these representations (Radvansky & Zacks, 2017). Consequently, it is important to demonstrate how this framework relates to the nature and construction of the mental representations people form of situations per se.

To this end, the following sections provide a focused and integrated review of three literatures in psychology that have each examined the nature of our mental representations of situations – social (and personality) psychology, engineering psychology, and narrative comprehension. Although they share a common interest in the representations of situations, each literature is concerned with different kinds of situations; research in narrative comprehension is largely concerned with how people understand situations described in stories. On the other hand, research in social and engineering psychology is concerned with how people understand the situations of which they are a part, with engineering psychology focusing on situations that arise in sociotechnical systems (e.g., an airplane cockpit or a nuclear power plant control room). Despite the fact that these literatures operate largely independently of each other, we have identified many points of agreement about the properties of our representations of situations that we believe coalesce within the framework of EST. In the following sections, we will discuss each point of agreement and in turn, demonstrate how EST embodies each property of our mental representations of situations.

1.1 Dualism and Correspondence

Each literature acknowledges a dualism comprising the external situation and the internal representation of the situation (i.e., there is a difference between what is "out there" and what is "in the head"). In social and personality psychology, for example, the study of situations and their effect(s) on people typically assumes either an "objective" view of situations, which emphasizes objectively quantifiable situational information (e.g., people, objects, or locations), or a "subjective" view of situations, which emphasizes the evaluative characteristics that people ascribe to the objective situation (e.g., "positive" or "adverse"; Furr & Funder, 2004; Rauthmann, Sherman, & Funder, 2015). Recently, Rauthmann and colleagues have advanced a model of situation perception that is compatible with both perspectives; perceivers arrive at a mental representation of a situation only after selecting, filtering, evaluating, and interpreting objectively measurable features of the external situation (e.g., the persons or objects present) via implicit (i.e., quick and effortless) and explicit (i.e., slow and deliberate) information processing (Rauthmann, 2012; Rauthmann et al., 2015; Rauthmann, 2016). The resulting mental representation reflects the psychologically meaningful or consequential characteristics of the external situation (e.g., its positivity or adversity).

Relatedly, research in engineering psychology is often aimed at an individual operator's understanding of a dynamically changing environment (i.e., their "situation awareness"; Durso & Sethumadhavan, 2008). Methods for assessing an operator's situation awareness in a dynamic environment (i.e., query-based techniques; Durso, Truitt, Hackworth, Crutchfield, & Manning, 1995; Endsley, 1990) assume that there is a

"… 'ground truth' against which its accuracy can be assessed (e.g. the objective state of the world or the objective unfolding of events that are predicted)," (Parasuraman, Sheridan, & Wickens, 2008, p. 144; van Winsen, Henriqson, Schuler, & Dekker, 2015).

Lastly, when understanding the unfolding events in a narrative text, readers form a non-linguistic representation of the situation described in the text (i.e., a "situation model"; Bransford, Barclay, & Franks, 1972; van Dijk & Kintsch, 1983; Glenberg, Meyer, & Lindem, 1987; Kintsch, 1988; Zwaan & Radvansky, 1998). As a mental representation, the situation model is distinct from but coordinated with the reader's "surface-level" representation of the specific words in a text and their syntactic relations as well as the "text base," which is a propositional representation of the semantic or conceptual information contained within the sentence.

Each literature also recognizes the close relationship between psychological outcomes (e.g., affect, cognition, or behavior) and the correspondence (or isomorphism) between the external situation and internal situation. For example, Dekker and Lutzhoft (2004) observed that virtually all theories of situation awareness (SA) agree that a high correspondence between the external situation and one's representation of it equates to good SA whereas a low correspondence equates to poor SA (e.g., Bedny & Meister, 1999; Endsley, 1995; Flach, 1996; Sarter & Woods, 1991; Smith & Hancock, 1995). When an operator "loses" his or her situation awareness of a safety-critical environment (e.g., in the driver's seat, an airplane cockpit, or an operating room), the operator becomes more vulnerable to error and performance failures, which can carry devastating consequences (Durso & Alexander, 2010; Durso & Sethumadhavan, 2008).

Similarly, theories of narrative comprehension argue that successfully understanding a narrative is equivalent to forming a coherent model of the situations described in the text (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998; Zwaan, Radvansky, Hilliard, & Curiel, 1998). Indeed, reading interventions that target situation model updating (e.g., teaching readers to explicitly monitor discontinuities in the features of situations) improves general reading comprehension ability relative to a standard reading comprehension curriculum (Wassenberg, Bos, de Koning, van der Schoot, 2015 but see Bohn-Gettler, 2014). Moreover, there have also been explicit attempts in engineering psychology to leverage situation model theory in dynamic environments (e.g., to facilitate information search in air traffic control, Durso, Johnson, & Crutchfield, 2010; Durso, Rawson, & Girotto, 2007).

1.1.1 Event Segmentation Theory

EST also posits that our representations of activity guide our cognition and behavior, the fitness of which depends on the correspondence between our representations and the external world. Specifically, EST proposes that our perception of ongoing activity is guided by event models in working memory that represent "what is happening now," (e.g., making breakfast or fluffing a pillow; Zacks et al., 2007, p. 274). Specifically, event models represent the entities involved in an event (e.g., objects and people), properties of those entities (e.g., their physical characteristics or goals), the spatial-temporal context within which an event occurs, and the structural relations between entities in an event as well as the causal or temporal relationship(s) among a chain of events (Radvansky & Zacks, 2011; Radvansky & Zacks, 2014).

Just as the correspondence between the external and internal situation matters for psychological outcomes, the degree of correspondence between ongoing activity and one's segmentation of it is also important. In the event segmentation literature, "correspondence" is operationalized as the extent to which an individual segments an activity in a normative, rather than idiosyncratic, manner (i.e., tending to identify the same boundaries that other observers identify). For example, the more normatively one segments a range of everyday activities, the better he or she is at recognizing and recalling the temporal order of events from the same activities (Zacks, Speer, Vettel, & Jacoby, 2006) or from different activities (Kurby & Zacks, 2011). Moreover, Bailey, Kurby, Giovannetti, and Zacks (2013) found that the ability to segment everyday activities (e.g., making breakfast) in a normative manner was also positively related to the ability to efficiently perform a different set of everyday activities (e.g., packing a backpack with school supplies), even after controlling for relevant cognitive variables such as working memory capacity, semantic memory, and script knowledge.

1.2 Situations are Bounded

That we recognize when the current situation ends and a new situation begins implies that situations can be demarcated, and thus bounded, in time. Indeed, the bounded nature of situations is reflected by the fact that we reason and talk about situations as coherent entities (Edwards & Templeton, 2005; Rauthmann et al., 2014). For example, people reliably sort descriptions of everyday situations (e.g., "Having a drink with some friends in a pub," or "Going to the bank,"; Forgas, 1976) into groups based on their perceived similarity to each other (e.g., Battistich & Thompson, 1980; Forgas, 1976;

Magnusson, 1971; Yang, Read, & Miller, 2006) as can professional air traffic controllers when given pictures of radar displays and flight strips depicting different air traffic situations (Neissen, Eyferth, & Bierwagen, 1999).

Moreover, we spontaneously form impressions of situations and describe them with adjectives such as "pleasurable," "productive," or "bad," (Edwards & Templeton, 2005). In fact, our vocabulary for describing situations appears to be structured such that we can systematically differentiate one situation from another. For example, Parrigon, Woo, Tay, and Wang (2017) performed the largest-to-date lexical analysis of the characteristics of situations, having participants rate how well 851 different adjectives described a variety of everyday situations. Parrigon et al. (2017) concluded that our vocabulary for describing situations varies along seven dimensions, with each dimension reflecting a different characteristic of situations - *Complexity* ("How complex or intricate is the situation?"), *Adversity* ("How difficult or depleting is the situation?"), *Positive Valence* ("How positively charged is the situation?"), *Typicality* ("How common and straightforward, or novel and ambiguous, is the situation?"), *Importance* ("How well will the situation lead to the attainment of one's goals?"), *Humor* ("How humorous or lighthearted is the situation?"), and *Negative Valence* ("How negatively charged is the situation?").

Lastly, like other kinds of bounded entities (e.g., objects; Rosch, 1999), situations can be described at different levels of abstraction; for example, an ottoman is a kind of chair, which in turn is a kind of furniture. Regarding situations, "Going to the dentist for a root canal treatment," and "Coming too late to a buffet and all the food is gone," are each a kind of "negative situation" (Rauthmann, 2016, p. 95). Indeed, there have been a number of efforts to create a taxonomy of situations in terms of the groups, clusters, or classes of situations (see Rauthmann, 2015 or Yang, Read, & Miller, 2009 for a review). For example, van Heck (1984, 1989) produced a prominent taxonomy of situation classes (Parrigon et al., 2017; Ten Berge & de Raad, 1999; Rauthmann et al., 2014), which comprises interpersonal conflict, joint working and information exchange, intimacy and interpersonal relations, recreation, traveling, rituals, sport, excesses, serving, and trading situations.

Research on the comprehension of described situations also suggests that we represent situations as bounded entities. The situation models that readers form of narrative text are presumably enclosed in a spatial-temporal framework, within which a single or a series of story events unfold (Radvansky & Dijkstra, 2007; Radvansky & Wyer, 1999). Different events occurring within the same spatial-temporal framework (e.g., while watching a movie in a movie theater) are represented as parts of the same situation. However, events that occur within different spatial-temporal frameworks are represented in different situations models (Radvansky & Zacks, 1991; Radvansky, Zwaan, Federico, & Franklin, 1998). During reading, for example, information appearing just prior to a large temporal discontinuity (e.g., "a day later.") is less accessible than when the same information appears before a small temporal discontinuity (e.g., "a moment later,"; Radvansky, Copeland, Berish, & Dijkstra, 2003; Zwaan, 1996). A large temporal discontinuity likely signifies the start of a new situation, requiring that the reader establish a new situation model, whereas a small discontinuity likely indicates a continuation of the same situation, requiring that the reader integrate the new event with the current situation model. Indeed, the cognitive effort required to establish a new situation model is reflected in longer reading times for sentences describing a large versus small temporal discontinuity. Similar results have obtained following spatial discontinuities (e.g., moving to a new and distinct location; Morrow, Greenspan, & Bower, 1987; Rinck & Bower, 2000; Rinck, Hähnel, Bower, & Glowalla, 1997; Zwaan et al., 1998), although reading time is not consistently affected as readers appear to be facile at updating along the spatial dimension (Radvansky & Copeland, 2010). Lastly, Radvansky (2005) argues that readers create a single situation model of multiple sentences if they all describe a common situation, but otherwise, create separate situation models of each sentence. For example, “The potted palm is in the hotel,” “The potted palm is in the museum,” and “The potted palm is in the barber shop,” refer to three different situations whereas, “The welcome mat is in the laundromat,” “The pay phone is in the laundromat,” and “The oak counter is in the laundromat,” all refer to a single common situation. After memorizing these sentences, verifying that “The potted palm is in the hotel,” is more difficult than “The welcome mat is in the laundromat,” because there are presumably three competing (i.e., interfering) situation models involving a potted palm but only one situation model involving a welcome mat.

1.2.1 Event Segmentation Theory

That we perceive boundaries between contiguous situations makes situations similar to the “events” examined in the event cognition literature. According to EST, events are “…a segment of time at a given location that is conceived by an observer to have a beginning and an end," (Zacks & Tversky, 2001, p. 17) and comprise goal-directed human activities lasting from seconds to tens of minutes. In this sense, situations are those events that extend farther in time than events such as basic actions (e.g., reaching or pushing) or more complex composite events (e.g., greeting a friend or ordering from a restaurant menu). However, situations are briefer than events like a baseball game or a wedding reception, which comprise a series of distinct, but related situations. Thus, situations fall somewhere between these extremes as "… momentary and fleeting phenomena that dynamically flow into each other," (Rauthmann & Sherman, 2016, p. 3) yet also have discernable boundaries in time and space.

To measure the segmentation of ongoing activity, Newtson (1973) devised a "unit marking" procedure in which participants watched a video of an actor performing a sequence of behaviors (e.g., filling out a questionnaire, lighting a cigarette, throwing out a match, and so forth). Before watching the video, participants were instructed to mark off the behavior of the actor into either the smallest units that seemed natural and meaningful (i.e., fine-grained segmentation) or into the largest units that seemed natural and meaningful (i.e., coarse-grained segmentation) by pressing a button. Newtson (1973) found that participants reliably varied the size of their units of perception, producing more fine-grained than coarse-grained units, and that consensual coarse-grained breakpoints tended to be a subset of consensual fine-grained breakpoints.

Later, Newtson and Engquist (1976) found that breakpoints identified in the unit marking procedure formed the boundaries of coherent perceptual units. Specifically, participants were better at detecting deletions of breakpoints than non-breakpoints from films, were more accurate at describing actions when viewing slides of only breakpoints versus slides of only non-breakpoints and had better recognition memory for breakpoints than non-breakpoints in films. More recently, Hard, Recchia, and Tversky (2011) found that participants looked longer at breakpoints than non-breakpoints when viewing slideshows of everyday activities at their own pace, and that looking time at coarse breakpoints, but not overall looking time, predicted the number of actions they could later recall.

1.3 Situations are Hierarchical

The boundaries between situations punctuate a dynamically unfolding series of events. Taking a similar view in social psychology, Forgas (1976) conceptualized situations as "social episodes" which he defined as "… interaction sequences which constitute natural units in the stream of behavior … [that] are distinguishable on the basis of symbolic, temporal, and often physical boundaries," (Forgas, 1976, p. 81). Indeed, when people recall and freely describe a situation they encountered recently, their descriptions often refer to episodes of activity; for example, " Playing chess," "Going to the bank," "Visiting your doctor," (Forgas, 1976, p. 85), "On a date with your boy/girlfriend," "Listening to a lecture in class," "Cleaning your room/house alone," (Battistich & Thompson, 1980, p. 77), "At the supermarket," "Going window shopping," "Having breakfast," (Eckes, 1995, p. 369), "Playing games at a friend's apartment," "Making dinner for me and my boyfriend," or "…watching TV," (Sherman, Nave, & Funder, 2010, p. 337). Thus, it seems natural for people to think of situations as episodes (i.e., a series of events bookended by situation boundaries).

In dynamic environments, Deutsch, Pew, Rogers, and Tenney (1994) observed that events, either externally driven or produced by an operator, occur within situations. Indeed, maintaining adequate SA requires that one recognize and integrate a series of events over time (Durso, Bleckley, & Dattel, 2007; Sarter & Woods, 1999; Woods, 1988), particularly when relying on SA to diagnose the cause of or prevent a problematic situation from arising. From several studies of highly skilled teams of NASA flight controllers, Christoffersen, Woods, and Blick (2007, p. 83) observed that "…event descriptions were found to be a highly prominent feature of the communication exchanged between flight controllers during monitoring, diagnosis, and replanning following anomalies," suggesting that flight controllers naturally conceive of system perform in terms of events. Moreover, Flach, Bennet, Jagacinski, Mulder, and van Passen (2004) observed that pilots describe poor SA as the inability to see how events cohere or connect to one another. Flach et al. (2004) argue that situations can be decomposed and described hierarchically at different levels of abstraction (e.g., from individual actions, such as a pilot adjusting the throttle, to the pilot's overall goal, such as landing safely). When the relations between the levels of abstraction are unclear, there is no longer a basis for decomposing activity into coherent chunks, rendering one's SA poor.

Finally, each literature agrees that our representations of situations are built from our knowledge of the world. Forgas (1982) argued that the representations of social episodes are rooted in the scripts or schemata that people form of the everyday, routine situations within their cultural milieu. Indeed, people across the adult lifespan largely agree on the basic sequence of actions that comprise routine situations (e.g., shopping at the grocery store or eating at a restaurant; Rosen, Caplan, Sheesley, Rodriguez, & Grafman, 2005) and how those basic actions are chunked to form larger meaningful units of activity (Bower, Black, & Turner, 1979). For example, most people agree that getting ready for work comprises turning off the alarm, waking up, getting out of bed, going to the bathroom, and so on (Rosen et al., 2005). Having structured representations of situations imparts a "… sense of coherence and stability and a hierarchical structure to the otherwise complex and confusing ebb and flow of social life," (Forgas, 1982, p. 68). These representations, in turn, have important consequences for cognition. For example, Flores, Bailey, Eisenberg, and Zacks (2017) observed that when people recall information from descriptions or depictions of routine activity, their responses are more likely to contain information that is directly related (rather than unrelated) to their pre-existing scripts for that activity and to organize their responses around (or distort the order of events to fit) their scripts (e.g., Abbott, Black, & Smith, 1985; Bartlett, 1932; Bower et al., 1979; Bower & Clark-Meyers, 1980; Brewer & Dupree, 1983; Lichtenstein & Brewer, 1980; Migueles & García-Bajos, 2012).

Lastly, situation models of dynamic environments are also believed to derive from "mental models," which are long-term memory structures that represent the causal and functional relations between the elements in a system (e.g., a nuclear power plant; Doane, Sohn, & Jodlowski, 2004; Durso et al., 2007; Durso & Gronlund, 1999; Endsley, 2000). A similar idea can be found in narrative comprehension; because narratives rarely describe situations completely nor cohesively, readers must often make inferences by drawing upon their knowledge to create a coherent situation model (Graesser, Singer, &

Trabasso, 1994). In this sense, situation models in narrative comprehension are amalgamations of information stated directly in the text and information summoned from the reader's knowledge (Zwaan & Radvansky, 1998); similarly, situation models of dynamic environments are amalgamations of information in the environment and the operator's knowledge base (Durso & Gronlund, 1999).

1.3.1 Event Segmentation Theory

That people perceive, conceive, and remember activity in a hierarchical fashion are among the key phenomena that EST strives to explain. EST claims that we maintain event models for different timescales simultaneously in working memory, although observers may selectively attend to event models for a specific timescale (e.g., fine or coarse-grained segmentation; Radvansky & Zacks, 2014). For example, when activity is coherent and hierarchically structured, observers can choose whether to attend to their fine- or coarse-grained event models (e.g., depending on their goals or task instructions, Zacks et al., 2007). However, when activity is less coherent and difficult to organize into coarse-grained units, perceivers may be forced to parse activity only at a fine-grained level. For example, observers tend to segment an activity more frequently when they cannot decipher an agent's goals (Hard, Tversky, & Lang, 2006; Newtson 1973; Wilder, 1978a; Wilder, 1978b) than when an agent's goal(s) are clear. Moreover, observers segment an activity less frequently as their familiarity with that activity increases, either through repeated exposure in the laboratory (Hard et al., 2006) or previous exposure within the domain from which the activity comes (e.g., expert-novice differences; Bläsing, 2015). Regarding the latter, professional dancers segmented a dance phrase into larger units than did non-dancers and moreover, amateur dancers segmented the same dance phrase into larger units after learning how to perform it than before learning it.

Zacks et al. (2001) were the first to demonstrate that within observers, events are perceived on multiple timescales such that fine-grained events are grouped within coarse-grained events. Zacks et al. (2001) had participants segment videos of everyday activities (e.g., making a bed once) twice, once under fine- and then under coarse-grained instructions (in a counterbalanced order). While segmenting, half of the participants also described each event that they perceived. Zacks et al. (2001) demonstrated hierarchical alignment of coarse and fine unit boundaries using both a discrete measure of alignment (based on co-occurrence of coarse and fine unit boundaries in 1-second time bins) as well as a continuous measure of alignment (based on the mean temporal distance between coarse and fine boundaries). Moreover, the hierarchical relationship between coarse and fine units was also apparent in how participants described events. For example, descriptions of coarse units tended to differ with respect to the object of interaction (e.g., a bed sheet versus a blanket) whereas descriptions of fine units within the same coarse unit differed with respect to the action performed on the object (e.g., unfolding versus straightening a bed sheet). Furthermore, of all the fine units within a coarse unit, descriptions of fine units nearest the boundaries of coarse units were most similar to the description of the overall coarse unit. The hierarchical alignment of coarse and fine unit boundaries within observers has since been replicated when segmenting videos of naturalistic activity (e.g., Kurby & Zacks, 2011; Sargent et al., 2013; Zacks, Kumar, Abrams, & Mehta, 2009; Zacks, Speer, Swallow, & Maley, 2010; Zacks, Speer, &

Reynolds, 2009; Zacks, Swallow, Vettel, & McAvoy, 2006), abstract geometric animations depicting intentional or randomly generated activity (e.g., Hard et al., 2006; Zacks, 2004), and narrative text (e.g., Kurby & Zacks, 2012; Zacks et al., 2009). In summary, the hierarchical structure of activity is a central and highly replicable characteristic of event segmentation both between and within observers.

Segmenting activity hierarchically is not merely something people can do when instructed, but rather, it appears to be a normal part of ongoing perception. In support, Zacks et al. (2001) used functional magnetic resonance imaging (fMRI) to scan the brains of participants while they watched short videos of everyday activities (e.g., making a bed; Zacks et al., 2001) three times each; during the first scan, participants passively watched the video without knowing that they would later segment it. During the second and third scans, participants pressed a button to segment the video into coarse or fine units. Zacks et al. (2001) found that brain regions (involved in attention-shifting and motion-processing) showed similar patterns of activity at event boundaries, regardless if someone was actively segmenting or passively watching the film. These findings have since been replicated using videos of simple geometric animations (Zacks, Swallow, Vettel, & McAvoy, 2006), extended naturalistic activity (i.e., the narrative film; Zacks et al., 2010), and narrative descriptions of everyday activity (Speer, Zacks, & Reynolds, 2007), which suggests that the segmentation of activity occurs automatically, regardless of whether activity is observed or described. Moreover, a common finding across these studies (c.f. Zacks et al., 2010) is that brain responses are generally larger at coarse- than fine-grained boundaries (Kurby & Zacks, 2018; Speer et al., 2007; Zacks et al., 2001; Zacks et al., 2006), suggesting that neural activity is modulated by the hierarchical structure of events (Zacks et al., 2001).

More recently, Baldassano et al. (2017) took a data-driven approach to identify event boundaries a priori as shifts between stable patterns of activity in regions throughout the brain. Baldassano et al. (2017) had subjects passively watch a 50-minute narrative film while having their brains scanned and later, applied a Hidden Markov Model to identify transitions, or boundaries, between stable patterns of cortical activity. Transitions separated by seconds in low-level sensory cortices (e.g., early and late visual areas) were hierarchically aligned with transitions separated by minutes in higher-order cognitive areas (e.g., those involved in episodic memory and tasks requiring the use of situation models; Ranganath & Ritchey, 2012). Lastly, Baldassano et al. (2017) also compared the boundaries extracted from the imaging data to the boundaries between "scenes" identified by a separate group of participants, finding that overt segmentation of the film aligned more closely with the coarser- than the finer-grained boundaries extracted from the imaging data.

EST posits that the hierarchical structure of event segmentation reflects the hierarchical structure of event knowledge (Hard et al., 2006). Event models in working memory are believed to be tokens of event schemata, which are structured representations in long-term memory of routine activities (e.g., making breakfast, eating at a restaurant, or doing the laundry; Shank & Abelson, 1977). While processing ongoing activity, event models predict the near future based on what has happened recently as well as information gleaned from event schemata (e.g., about how a sequence of events might unfold).

Studies that have allowed people to segment at a preferred or natural grain have consistently found knowledge effects despite the fact that knowledge has been manipulated in different ways, such as having knowledge about an actor's goals (Newberry & Bailey, 2019; Newtson, 1973; Wilder, 1978a, Wilder, 1978b), comparing domain experts with novices (Bläsing, 2015, Levine, Hirsh-Pasek, Pace, & Golinkoff, 2017; Markus, Smith, & Moreland, 1985), or having prior knowledge about an actor or events (Graziano, Moore, & Collins, 1988; Massad, Hubbard, & Newtson, 1979). A common result across studies using "natural" segmentation instructions is that having knowledge of some kind enables one to parse activity into larger units, relative to having no or less knowledge.

However, evidence that event knowledge is related to the hierarchical structure of online event segmentation is somewhat inconsistent. For example, Zacks et al. (2001) found that the hierarchical alignment of coarse and fine event boundaries increased the more familiar an observer was with an activity (e.g., assembling a saxophone) and also when an observer described the activity while segmenting it, which presumably encouraged the use of knowledge in justifying one's decisions. Neither of these results, however, have since replicated (Hard et al., 2006; Zacks & Kurby, 2011). Moreover, Sargent et al. (2013) found that a measure of script knowledge was positively related to the normativeness of one's segmentation, but not significantly so. Thus, knowledge clearly influences online event segmentation, but its effect(s) is difficult to detect when observers must explicitly segment activity into small and large units (i.e., hierarchically, but see Swallow, Kemp, & Simsek, 2018).

1.4 Situations are Multifaceted

Across the literatures, there is consensus that our representations of situations are multifaceted and moreover, there is convergence in terms of what those facets are. Research on situations as experienced (i.e., in social and personality psychology) and situations as described or depicted (i.e., in narrative comprehension) appear to agree on the basic features of our representations of situations; Who or what is present, what is happening, when and where is the situation happening, and why is this happening (Johns, 2006; Parrigon et al. 2017; Rauthmann, Sherman, & Funder, 2015; Saucier, Bel-Bahar, & Fernandez, 2007)? These categories of situational information have collectively been called “situation cues” (Rauthmann, 2015) and are known to constitute the mental representations people form of described situations (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). When these features of situations change in a story, readers tend to update their current situation model to reflect the altered situation, which is evidenced by increases in processing time (i.e., reading time; Bailey, Kurby, Sargent, & Zacks, 2017; Rinck & Weber, 2003; Zacks et al., 2009; Zwaan, Magliano, Graesser, 1995; Zwaan, Radvansky, Hilliard, & Curiel, 1998) as well as reduced accessibility of information that appeared just prior to a change (Bailey et al., 2017; Ditman, Holcomb, & Kuperberg, 2008; Glenberg et al., 1987; Morrow, Greenspan, & Bower, 1987; Radvansky & Copeland, 2010; Rinck & Bower, 2000; Speer & Zacks, 2005; Zwaan, 1996; Zwaan, Madden, & Whitten, 2000). Outside of narrative text, changes in situation cues in virtual reality environments (e.g., spatial location; Radvansky & Copeland, 2006; Radvansky, Tamplin, & Krawietz, 2010) and narrative films also produce situation model updating effects (Huff, Meitz, & Papenmeier, 2014; Magliano, Miller, & Zwaan, 2001), suggesting that models of narrative comprehension may be generalizable to the mental representations of everyday experiences (Radvansky & Zacks, 2011).

However, researchers in social and personality psychology (e.g., Edwards & Templeton, 2005) recognize that our representations of situations also reflect the "characteristics" of a situation (e.g., positive, social, effortful, or sexual; Rauthmann, 2015; Rauthmann et al., 2015), which are believed to be interpretations or evaluations of the meaning of a constellation of situation cues (i.e., the gestalt; Pervin, 1978). For example, situations perceived as "positive" are more likely to contain friends [Who?] than situations containing working or studying [What?] (Rauthmann et al., 2014). Furthermore, two people can be exposed to the same set of situation cues but differ in the specific qualities they ascribe to the overall situation (e.g., is the situation threatening or is it enjoyable?), which can arise from how observers differentially attend to and evaluate situation cues (e.g., due to different personality traits, moods, or goals; Rauthmann et al., 2015). Of note, Rauthmann and Sherman (2018) observed that five replicable situation characteristics have emerged across recent large-scale independent research efforts; these characteristics describe the extent to which a situation affords or requires either 1) overcoming external threats and obstacles, 2) dealing with internal negative events that may cause distress, 3) getting an important or urgent task accomplished, 4) using deeper and effortful cognitive processing, or 5) engaging with fun and pleasant events.

In engineering psychology, several researchers have lamented a lack of understanding about what differentiates one situation from another (Flach, 1995; Flach et al., 2004; Tenney & Pew, 2006). Indeed, the kinds of information reflected in an operator's model of a situation tend to be domain-specific (e.g., a patient's vital signs are relevant to an anesthesiologist whereas altitude and heading are relevant to a pilot), making generalization across situations difficult. However, Durso et al. (2007) argued that models of narrative comprehension, which have categorized the situational information that readers routinely monitor, may also serve as a model of situation awareness in dynamic environments. In particular, the Event-Indexing Model (Zwaan & Radvansky, 1998) posits that situation models primarily represent the events described in a narrative (Radvansky & Copeland, 2010; Rinck & Weber, 2003; Zwann, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998), such as a "… a raindrop falling, an explosion, or someone tripping over a cat," (Zwaan, 1999, p. 95). For each event, readers construct an index for when the event occurred (e.g., a moment or a day later), where in space it occurred, why it occurred (i.e., the physical or motivational cause of the event), and who or what was involved in the event (e.g., characters or objects). Indeed, Durso et al. (2010) found that information sought by operators monitoring dynamic real-word situations overlaps with the information that readers naturally monitor when reading narratives. Durso et al. (2010) classified the information requests of air traffic controllers with respect to the five dimensions of situational continuity identified by the Event-Indexing Model (Zwaan & Radvansky, 1998). Specifically, traffic information requests related to protagonists (i.e., characteristics of an aircraft), intentionality (i.e., plans of an aircraft), and space (i.e., location of an aircraft) accounted for the majority of information requests. Moreover, controllers preferred information displays that were organized around these dimensions (i.e., when pieces of information related to the same dimension appeared next to each other on the display) more than those that were not (i.e., when the same information was interdigitated).

1.4.1 Event Segmentation Theory

According to EST, event segmentation depends upon the processing of feature changes (Zacks, Kumar, Abrams, & Mehta, 2009), which include physical as well as conceptual changes in a situation. Regarding the physical features of activity, event boundaries coincide with changes in motion parameters (e.g., speed or direction), especially fine-grained event boundaries (Hard et al., 2006; Newtson, Engquist, & Bois, 1977; Zacks, 2004; Zacks et al., 2009). Regarding the conceptual features of activity in narrative texts and films, changes in each of the five dimensions of situational continuity identified by research on narrative comprehension (i.e., the Event Indexing Model; Zwaan & Radvansky, 1998) reliably lead to the perception of event boundaries at both the fine and coarse-grained levels of activity: time (Kurby & Zacks, 2012; Speer & Zacks, 2005; Speer et al., 2007; Zacks et al., 2010), space (Bailey et al., 2017; Kurby, Asiala, & Mills, 2014; Kurby & Zacks, 2012; Speer et al., 2007; Zacks et al., 2009; Zacks et al., 2010), characters (Bailey et al., 2017; Kurby et al., 2014; Kurby & Zacks, 2012; Speer et al., 2007; Zacks et al., 2009; Zacks et al., 2010), causality (Kurby et al., 2014; Kurby & Zacks, 2012; Speer et al., 2007; Zacks et al., 2010), and intentionality (Kurby et al., 2014; Kurby & Zacks, 2012; Speer et al., 2007; Zacks et al., 2009; Zacks et al., 2010).

Similar findings obtain when observers indicate each time the situation or circumstances in a narrative story or film changes, although these segmentation instructions are relatively rare in the literature. Magliano, Kopp, McNerney, Radvansky, and Zacks (2012) instructed participants to indicate each time they perceived that the "overall situation" had changed in either text or picture-only versions of children's stories. In both modalities, situation-change judgments coincided with the introduction of new characters, initiating events (i.e., events that motivate the beginning of a goal-directed action sequence), and the beginning and outcome of goal-directed action sequences. In narrative film, situation-change judgments (i.e., the points in the film in which the circumstances or situation changed) correspond to the beginnings and ends of goal-directed action sequences (Magliano, Taylor, & Kim, 2005) and changes in spatial-temporal framework (Magliano et al., 2001). Moreover, similar results obtain when people segment situations in non-narrative, loosely structured environments. For example, Magliano, Radvansky, Forsythe, and Copeland (2014) had participants segment a pre-recorded playback of a first-person shooter video game by indicating each time the situation changed for the player in the game. Magliano et al. (2014) found that the boundaries between situations corresponded to changes in space, the introduction of enemies, and the initiation or accomplishment of the player's superordinate goals (e.g., kill all enemies present).

Lastly, Mumma and Durso (in revision) found that changes in situation cues and situation characteristics (i.e., the "Situational Eight DIAMONDS"; Rauthmann et al., 2014) independently predicted the likelihood of perceiving the start of a new situation, fine-grained, or coarse-grained event in a narrative film (reproduced in Appendix C). Briefly, the DIAMONDS dimensions reflect the different ways in which people perceive, describe, and evaluate everyday situations and comprise *Duty* (e.g., “Does work need to be done?”), *Intellect* (e.g., “Is deep cognitive information processing relevant?”), *Adversity* (e.g., “Is someone under threat?), *Mating* (e.g., “Is the situation erotically charged?”), *pOsitivity* (e.g., “Is the situation enjoyable?”), *Negativity* (e.g., “Could the situation turn negative?”), *Deception* (e.g., “Is mistrust an issue?”), and *Sociality* (e.g., “Is meaningful social interaction and relationship building possible?”).

According to EST, changes in the features of situations coincide with event boundaries because these are moments when predicting the near future is most difficult. Event models generate “perceptual predictions,” or more accurately, representations of the state of the world in the near future (Zacks et al., 2007), that are continually compared with sensory input. Event models remain unchanged in the face of varying sensory input until an error-detection monitoring system recognizes that prediction error has become too large. When this occurs, the event model is replaced with a new model by conferring with available sensory input and event schemata. Once the new event model generates accurate predictions, it is again shielded from sensory input, remaining stable until the next spike in prediction error occurs. The process of updating an event model is subjectively perceived as a boundary between events and it is this moment that people report during overt segmentation tasks. In sum, EST claims that the segmentation of activity is a side effect of our perceptual system attempting to accurately predict the near future (Kurby & Zacks, 2008), which is more difficult when a situation is changing than when it remains constant.

EST offers one explanation of situation model updating effects during narrative comprehension; activity becomes difficult to predict when a situation changes in a story, and consequently, the reader must update his or her event model (Radvansky & Zacks, 2014). Indeed, the cognitive consequences of updating an event model are similar to those observed when readers are presumably updating their situation model. For example, information appearing shortly before an event boundary is less accessible than if there was no event boundary (Bailey, Kurby, Sargent & Zacks, 2017; Speer & Zacks, 2005) and sentences containing an event boundary are read more slowly than sentences that do not contain a boundary (Bailey et al., 2017; Speer & Zacks, 2005; Zacks et al., 2009). Consequently, some researchers treat the “situation models” of narrative comprehension as a kind of event model (i.e., one that is derived from language rather than directly from experience; Radvansky & Zacks, 2014; Zacks et al., 2009).

1.5 Situations are Future-Oriented

Across the literatures, our representations of situations are thought to play an important role in facilitating predictions of the future, which help one plan their behavior in a situation and respond to circumstances before they arise. However, this role appears to be far more central in our representations of experienced situations than in those of described situations. Briefly, situation models of narrative text do support predictive inferences, which correspond to future states of the world (e.g., the outcome of an event; Cook, Limber, & O'Brien, 2001; McDaniel, Schmalhofer, & Keefe, 2001; Peracchi &

O'Brien, 2004; Schmalhofer, McDaniel, & Keefe, 2002). During comprehension, however, these inferences are generated far less often than other kinds of inferences, particularly those that help achieve explanatory coherence (e.g., inferences about causality or the goals of characters that explain why events happen; Graesser et al., 1994; McNamara & Magliano, 2009; van den Broek, Beker, & Oudega, 2015). When readers do generate predictive inferences during reading, it is only when those inferences are highly constrained or supported by the preceding narrative context (Magliano, Dijkstra, & Zwaan, 1996). When watching a narrative film, however, viewers rely on inferences about the past and future to understand events (Levin & Baker, 2017). Unlike narrative text, narrative films depict richly situated contexts and moreover, filmmakers routinely use cinematic devices that encourage viewers to make specific predictions (e.g., via editing or framing; Magliano et al., 1996), further engaging the audience with the filmic world. The divergence between the literatures on this point should not be entirely surprising; predictive inferences are far more important in real (i.e., experienced) situations where it is highly advantageous to anticipate, rather than merely react, to what happens next in the situation.

In social psychology, the characteristics of situations are thought to reflect the underlying causal forces that operate in a situation. For example, if a situation is perceived as “difficult” or “boring,” then it suggests that the situation itself is exerting some kind of observable effect (e.g., on the behavior of the people in the situation; Edwards & Templeton, 2005). Recent taxonomies of situation characteristics (Parrigon et al., 2017; Sherman, Rauthmann, Brown, Serfass, & Jones, 2015; Rauthmann et al., 2014) have predicted the behaviors of people in situations from the situation's characteristics in conceptually meaningful ways (e.g., situations perceived to be "important" tend to elicit conscientious behavior; Parrigon et al. 2017). The ability to summarize a constellation of situation cues in terms of a smaller set of characteristics means that "…a perceiver is able to understand … what is happening, surmise what might have led to the observed state of affairs, extrapolate what might happen and coordinate [one's] own behaviour accordingly," (Rauthmann et al., 2015, p. 373).

In engineering psychology, theoretical models of situation awareness generally follow either an information processing framework (e.g., Endsley, 1995) or are explicitly adapted from Neisser's (1976) perceptual cycle (e.g., Adams, Tenney, & Pew, 1995; Smith & Hancock, 1995). However, common to all models is the idea that our representations of situations generate expectations (e.g., about the values of system parameters) and that an unexpected change (Flach, Fuefel, Reynolds, Parker, Henrickson, & Kellogg, 2017; Landman, Groen, van Paasen, Bronkhorst, & Mulder, 2017) or the absence of an expected change (Christoffersen et al., 2007) may suggest a poorly calibrated internal model. Endlsey (1995, p. 57) argues that "… the main clue to erroneous SA [situation awareness] will occur when a person perceives some new piece of data that does not fit with expectations based on his or her internal model … A common problem is whether to continue to revise the existing model to account for the new data or choose an alternate model that is more appropriate. For the latter to occur, something about the data must flag that a different situation is present." Moreover, Wickens (2015) observes that in many dynamic systems, understanding and responding to the current state of the system is sometimes less important than responding to the future state, particularly in systems in which responses take time to execute (e.g., moving troops to a new location). The ability to "stay ahead" of the situation has been found to be predictive of performance in a range of dynamic systems (e.g., Engström et al., 2018; Ma & Kaber, 2007; O'Brien & O'Hare, 2007; Sulistyawati, Wickens, & Chui, 2011).

1.5.1 Event Segmentation Theory

Prediction is *the* heart of EST because anticipating the future accurately enables organisms to plan their behavior (Eisenberg, Zacks, & Flores, 2018; Zacks et al., 2007). Support for the role of prediction in event segmentation comes from both computational and behavioral evidence. Regarding the former, Reynolds, Zacks, and Braver (2007) implemented the architecture of EST in a neural network. After training, the model was able to predict the motions of a 3-dimensional figure performing routine tasks (e.g., chopping down a tree) by using spikes in prediction error to identify event boundaries and update its model of the current event accordingly. Regarding behavioral evidence, Zacks, Kurby, Eisenberg, and Haroutunian (2011) had participants watch videos of everyday activities (e.g., washing a car or pitching a tent), during which the video was stopped 2.5 seconds before either an event boundary or an event middle (as determined by normative segmentation data). Each time the video stopped, participants rated their confidence in their ability to predict what would happen in the near future. Then, participants chose which of two video-stills (i.e., the actual frame or a foil) would most likely occur 5 seconds later. A subsequent study repeated these procedures using a Yes/No decision task rather than a two-alternative forced choice task. In both studies, participants were slower, less accurate, and less confident when making predictions before an event boundary than when making predictions before the middle of an event. Rather than overtly assessing the relationship between predictability and event structure, Eisenberg et al. (2018) covertly measured viewers expectations while segmenting videos of everyday activities in terms of their “predictive looking” behavior (i.e., anticipatory eye-movements towards objects that would eventually be contacted by an actor). Eisenberg et al. (2018) measured the amount of time participants spent looking at a target object 3 seconds before an actor touched it. When the target object was contacted within an event, participants looked ahead to the target object earlier than when the target object was contacted near an event boundary.

Again, EST claims that event boundaries coincide with situational changes because it is harder to predict the future at these moments than when the situation is relatively constant. Thus, predictive accuracy and the likelihood of perceiving an event boundary should each be related to how much the situation is changing at a given moment. Indeed, as the number of situational changes in a clause increase (e.g., changes in characters or spatial location), the perceived predictability of that clause decreases while the probability of identifying an event boundary increases (Zacks et al., 2009); and, predictive accuracy decreases, and response latencies increase, as the number of situational changes at an event boundary increases (i.e., characters, time, location, and action; Huff et al., 2014).

1.6 Present Studies

In summary, our mental representations of situations are bounded, hierarchical, multidimensional, and future-oriented (at least for experienced situations). However, the important question remains as to how our representations of situations, regardless of where they exist (i.e., in the world, text, or film), are formed and modified. The present studies argue that EST provides a theoretical framework that is qualified to supply an answer. Because Mumma and Durso (in revision; reproduced in Appendix C), have already demonstrated that changes in the facets of situations (i.e., situation cues and characteristics) increase the likelihood of perceiving a boundary between situations (as well as fine-grained events), the primary goal of the present studies is to understand the cognitive mechanisms that give rise to the hierarchical structure of our representations of situations.

According to EST, an event boundary identified during overt segmentation corresponds to an event model being updated in response to prediction error, with fine boundaries marking the advent of a new fine-event model and coarse boundaries marking the advent of a new coarse-event model. Although coarse-event models are rarely updated without also updating a fine-event model (Radvansky & Zacks, 2011), fine-event models are often updated during coarse events (i.e., without updating the coarse-event model). In the latter case, EST predicts that whenever a fine event boundary occurs within the boundaries of a coarse event, the coarse event model should experience less prediction error at that moment than the fine event model.

As an illustration, consider the example of going to a coffee shop (Zacks & Sargent, 2010). If one is attending to activity at the coarse level of the situation (i.e., going to a coffee shop), then one expects that a set of certain fine events will occur (e.g., entering the coffee shop, waiting in line, ordering at the register, waiting for your order, receiving your order, and leaving the coffee the shop). The beginning of these fine events coincides with changes in the situation, such as a change in location (e.g., when moving from the door to the end of the line), the introduction of a new person (e.g., a barista when ordering), or the start of a new action (e.g., reaching for a cup). Each of these changes are inconsistent with the expectations of the current fine event model (e.g., what happens when ordering at a register is quite dissimilar to what happens when waiting in a line) and consequently, prediction error increases to the point that a new fine model must be instantiated. For the event model of the overall situation, however, these changes should not increase prediction error because they are each consistent with what one expects should happen in the situation. It is not until one begins to the leave the coffee shop that the model of the situation no longer generates accurate predictions and consequently, a more appropriate model of both the situation and current fine event must be instantiated (e.g., walking down the street).

However, the hypothesis that differential predictive accuracy underlies the hierarchical structure of event models has not been tested; we know that predictive accuracy, measured either overtly (e.g., making Yes/No or two-alternative forced decisions; Huff et al., 2014; Zacks et al., 2011) or covertly (e.g., making predictive eye-movements; Eisenberg et al., 2018) is better when within an event than when approaching the end of an event (see also Reynolds et al., 2007). Critically, we do not know whether predictive accuracy differs *between* event models when a fine-event model is updated but the superordinate (i.e., coarse) event model is not.

To this end, we conducted a pair of studies in which participants segmented a narrative film (*The Red Balloon*; Lamorisse, 1956) either into situations (i.e., each time a new situation has begun), fine events (i.e., each time a new small meaningful unit of activity has begun) or events at a neutral granularity (i.e., each time a new meaningful unit of activity has begun). This manipulation oriented participants to event models on a specific timescale (Zacks et al., 2007) and also served as a manipulation check. The film stopped 18 times, each time in one of three places that were determined from previously collected normative segmentation data: 1) within an ongoing situation and fine event (Within/Within), 2) when nearing the end of a fine event in an ongoing situation (Within/Across), or 3) when nearing the end of both a fine event and situation (Across/Across). Each time the film stopped, participants rated their confidence in their ability to predict what would happen next in the film (Zacks et al., 2011), which reflects the state of the error-detection monitoring system (Richmond & Zacks, 2017). Afterwards, we assessed their predictive accuracy and participants then resumed segmenting the film from where it stopped.

We examined the effect of hierarchical event structure on predictive accuracy by following the “three-pronged method” for studying inference generation (Magliano & Graesser, 1991). Briefly, the first prong is to use a theoretical model to predict the kind(s) of inferences that should be produced during comprehension. In our case, EST proposes that people generate predictive inferences during the comprehension of ongoing activity. The second prong is to expose inferences by using a think-aloud or question-answering procedure during comprehension (e.g., answering “What happens next?” at different moments in a passage; Graesser et al., 1994). In the first study, participants were asked to write down their prediction of what would happen next in the film at each stopping point and later, a rating of the consistency of their prediction with what happened in the ensuing 5 seconds was obtained. The third prong is to use online behavioral measures (e.g., a recognition or lexical decision task) to demonstrate convergence with the verbal inference elicitation procedure. In the second study, participants were given either a description of what happened next in the film (i.e., a target) or a foil (i.e., an uncommon and incorrect prediction) at each stopping point and then decided whether that would happen next in the film.

Assessing predictive accuracy using both a question-answering procedure and a Yes/No decision task ensured that our conclusions would not be endemic to one approach. For example, previous studies have assessed the effect of event structure on predictive accuracy via two-alternative forced decision (Huff et al., 2014; Zacks et al., 2011) or Yes/No decision tasks (Zacks et al., 2011). These tasks have the disadvantage that participants can answer them by performing a simple feature comparison between the probe (i.e., a target or foil) and the preceding context (Magliano & Graesser, 1991), rather than a prediction per se. Answering what-happens-next questions avoids this problem, but unlike decision tasks, forces participants to verbalize their perceptual predictions, which may be difficult to articulate. Thus, convergence across both studies would only strengthen our conclusions whereas divergence would, at the very least, be instructive. Regarding the primary goal of the present studies, we hypothesized the following:

Hypothesis 1: Predictive accuracy and confidence should be greater when within an event than when approaching an event boundary, but only when within an event or near an event boundary on the timescale to which one is oriented (i.e., to situations or fine events; Figure 1).

Hypothesis 2: The predictive accuracy of and confidence in event models of situations should only be greater than that of fine-event models when nearing the end of a fine event in an ongoing situation (i.e., Within/Across; Figure 1).

Because the following studies were the first to measure subjective confidence, predictive accuracy, and segmentation, we also tested several critical, but unexamined hypotheses regarding the general mechanisms of EST:

Hypothesis 3: Confidence in one's ability to predict what happens next should be negatively related to the likelihood of perceiving an event boundary in the near future.

Hypothesis 4: The accuracy of one's prediction about what happens next should be negatively related to the likelihood of perceiving an event boundary in the near future.

Hypothesis 5: Confidence in one's ability to predict what happens next should be positively related to the accuracy of one's prediction.

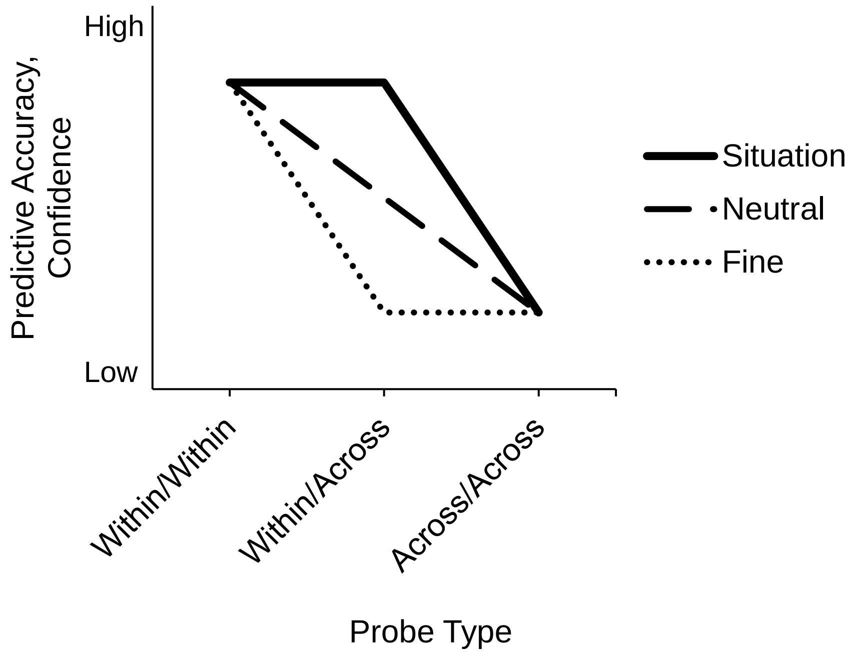

Figure 1 – The hypothesized effects of Orientation (Situation, Neutral, Fine) and Probe Type (Within/Within, Within/Across, Across/Across) on predictive accuracy and confidence in Experiments 1 and 2.

CHAPTER 2. EXPERIMENT 1

Each time the film stopped in Experiment 1, participants rated their confidence in their ability to predict what would happen next and then wrote down their prediction of what would happen next in the film. This approach is consistent with the assumption that event models are working memory representations and consequently, their contents (i.e., predictions) should be reportable while comprehending ongoing activity (Hard et al., 2006; Kurby & Zacks, 2011; Kurby & Zacks, 2012; Zacks et al., 2001; Zacks et al., 2007). According to EST, prediction error is associated with subjective uncertainty (Richmond & Zacks, 2017; Zacks et al., 2011) and should be minimal as long as sensory input is consistent with the prediction(s) of an event model (Huff et al., 2014; Zacks et al., 2007). Accordingly, the accuracy of each prediction was rated in terms of its consistency with what happened next in the film (i.e., in the 5 seconds of the film immediately following a probe).

2.1 Method

2.1.1 Participants

We targeted a sample size of 90 participants (30 in each level of Orientation) to detect a moderately sized interaction (Cohen's $f = .25$; Cohen, 1972) between Probe Type and Orientation with 80% power. Ninety-five undergraduate students from the Georgia Institute of Technology participated in partial fulfillment of a research familiarization requirement. All data for five participants were excluded because they reported on the post-experiment questionnaire having seen the film before. These participants were replaced. The remaining 90 participants had a mean age of 19.6 years (*SD* = 1.26) and 56% were male.

2.1.2 Design

Experiment 1 used a 3 (Orientation: Situation, Neutral, Fine) x 3 (Probe Type: Within/Within, Within/Across, Across/Across) mixed design with Orientation varying between subjects and Probe Type varying within subjects. For each level of Probe Type within a participant, we measured his or her confidence in the ability to predict what would happen next in the film as well as how consistent his or her prediction was with what actually happened next. We also recorded the exact moment in time that participants reported an event boundary during the film.

2.1.3 Materials

2.1.3.1 Probe Locations

Using existing segmentation data of *The Red Balloon* (Kurby et al., 2014; Mumma & Durso, in revision), a separate normative situation and fine segmentation profile was created by dividing the film into 394 5-second time bins and then computing the probability of someone identifying at least one event boundary in each time bin. Time bins with a probability greater than or equal to .50 were considered consensual boundaries. In total, 67 consensual fine boundaries and 16 consensual situation boundaries were identified, 10 of which (63%) were also consensual fine boundaries. Thus, consensual situation boundaries tend to be a subset of consensual fine boundaries, just as coarse boundaries also tend to be a subset of fine boundaries, regardless of whether segmentation grain is manipulated between subjects (e.g., Hanson & Hirst, 1989; Newtson, 1973) or within subjects (e.g., Hard et al., 2006; Zacks et al., 2001). Moreover, the observed overlap between fine and situation boundaries is not attributable to chance; the joint probability of a time bin being both a consensual fine and situation boundary was (67/394) x (16/394) = .007. Using an exact goodness of fit test, the observed number of time bins containing both a consensual fine and situation boundary (10) was significantly greater than the number expected by chance (.007 x 394 = 2.76), $p = .001$.

Using the consensual situation and fine boundaries, we chose six probe locations that were within an ongoing situation and fine event (Within/Within), six when nearing the end of a fine event in an ongoing situation (Within/Across), and six when nearing the end of both a fine event and a situation (Across/Across). Within/Within probes were chosen from among the time bins with the lowest probability of being perceived as a fine and situation boundary, Within/Across probes from bins with the greatest probability of being perceived as a fine boundary but not a situation boundary, and Across/Across probes from bins with the greatest probability of being perceived as a fine and a situation boundary (Figure 2; Table 1). Additionally, probes were selected so that the frequency of each type of probe occurred as equally as possible over the four film clips (Table 2).

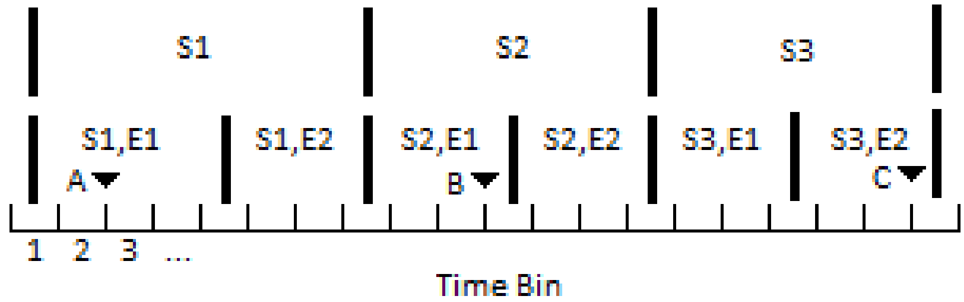

Figure 2 – Boundaries in the upper row demarcate situations (e.g., S1) and boundaries in the lower row demarcate fine events within a situation (e.g., S1, E1). Marker "A" is a "Within/Within" probe, "B" is a "Within/Across" probe, and "C" is an "Across/Across" probe.

Table 1 – Probability of perceiving a fine or situation boundary for each level of Probe Type.

	Within/Within		Within/Across		Across/Across	
	M	*SD*	*M*	*SD*	*M*	*SD*
Fine Boundary	.03	.04	.60	.06	.72	.12
Situation Boundary	.04	.02	.15	.04	.61	.09

Table 2 – Frequency of each type of probe in each clip.

Clip	Within/Within	Within/Across	Across/Across	Total Number of Probes
1	1	2	1	4
2	2	1	2	5
3	1	2	2	5
4	2	1	1	4

To determine when, precisely, each probe should occur in the film, the average (mean) timestamp for each of the chosen boundaries was obtained. Consistent with the methods of similar studies (Huff et al., 2014; Zacks et al., 2011), experimental probes involving a single event boundary (i.e., Within/Across) occurred 2.5 seconds before the average timestamp for that boundary. Probes involving two consensual boundaries (i.e., Across/Across) occurred 2.5 seconds before the average of the two average timestamps. Lastly, probes that did not involve an event boundary (i.e., Within/Within) occurred 2.5 seconds before the middle of a time bin designated as a Within/Within probe location. On average, probes occurred 1.88 minutes apart from each other (*SD* = 0.84).

2.1.3.2 Consistency Ratings

Three raters independently rated the consistency of all 1,620 predictions (90 participants x 18 predictions per participant) while blind to Probe Type and Orientation. Raters compared each prediction to a video clip of the 5 seconds that followed the corresponding probe location in the film. For each prediction, raters answered the question, "How consistent is the prediction with what happens next in the film?" using a scale that ranged from 1, "Not at all consistent," to 4, "Extremely consistent," (Table 3). For each participant under each rater, the mean consistency rating of the six Within/Within, Within/Across, and Across/Across probes was computed and subsequently averaged across raters for hypothesis testing. The interrater reliability of the average consistency rating for each level of Probe Type was assessed with three separate intraclass correlation coefficients (ICCs) using an average measures absolute agreement definition (ICC [2, k]; Shrout & Fleiss, 1979).

Table 3 – Consistency rating scale.

Value	Label	Description
1	Not at all consistent	What happens next completely precludes prediction.
2	Slightly consistent	What happens next likely precludes prediction.
3	Moderately consistent	What happens next does not likely preclude prediction.
4	Extremely consistent	Prediction describes something that happens next.

2.1.4 Procedure

After providing informed consent, participants completed a brief demographics questionnaire and were then randomly assigned to one of the three levels of Orientation (i.e., Situation, Neutral, or Fine). Next, participants were seated approximately 61 centimeters away from an 81-centimeter diagonal computer screen. The experimenter told participants that they would watch a film and (depending on their assigned level of Orientation) press the spacebar each time they felt that:

> Situation Orientation: a *new* situation has begun in the film. A “situation” is whatever you think the current circumstances or state of affairs in the film is.
>
> Neutral Orientation: a *new* meaningful unit of activity has begun in the film. What counts as a “meaningful unit of activity” is entirely up to you.
>
> Fine Orientation: a *new small* meaningful unit of activity has begun in the film. What counts as a “*small* meaningful unit of activity” is entirely up to you.

Lastly, the experimenter explained that at different times throughout the film, participants will also be asked to rate a statement and then write down what they think would happen next in the film.

All stimuli were presented with PsychoPy (Pierce, 2007). Participants practiced the experimental procedure while watching a 4.4-minute-long clip from *North by Northwest* (Hitchcock, 1959), which was stylistically similar to *The Red Balloon* (e.g., contained little dialogue and unfolded nearly continuously in space and time). At the end of the practice clip, participants rated their confidence in their ability to predict what would happen next in the film using a scale from 1, "Not at all," to 6, "Extremely," (Zacks et al., 2011). Immediately afterward, participants wrote down what they thought would happen next in the film in a booklet. When finished, they pressed the spacebar to resume the film and the program reminded them briefly of their assigned segmentation instructions. The experimenter remained in the room for the entire practice session to ensure that participants understood the task and to address any questions or concerns related to the task.

The participants then repeated these procedures while watching *The Red Balloon*, which lasted approximately 33 minutes and was divided into four successive clips, lasting 7.7, 7.8, 7.4, and 10 minutes, respectively (Zacks et al., 2009). Participants were allowed to take a brief break after finishing each clip. During each clip, the film stopped at the predetermined probe locations, participants rated their confidence and then wrote down a prediction in a small booklet with a single blank page for each probe. Afterwards, the program reminded the participant of their segmentation instructions and the film resumed playing from where it stopped. The booklets were replaced after each clip and always contained more cards than probes so that participants could not anticipate the number of probes in a clip. The experimental program recorded the participant's confidence rating at each probe location as well as the exact moment(s) he or she pressed the spacebar in between successive probe locations, which was rounded to the nearest whole second. At the end of the experiment, participants completed a post-experiment questionnaire about their experience and indicated whether they had seen *The Red Balloon* before.

2.1.5 Analyses

For all analyses, the criterion for significance was an α level of .05. When corrections for multiple comparisons were used (e.g., the Bonferroni correction), the reported *p* values are adjusted and can be compared to .05. The violation of sphericity in mixed model ANOVAs was assessed with Mauchly's sphericity test and if violated, degrees of freedom were adjusted using the Greenhouse-Geisser correction. As estimates of effect size, we report partial eta-squared ($\eta^2_{partial}$) for ANOVA effects and Cohen's *d* for *t*-tests.

2.2 Results

2.2.1 Manipulation Check

To determine whether the Orientation manipulation successfully altered how participants segmented the film, we examined the degree to which situation and fine segmentation in the present study were related to normative segmentation data from previous studies using *The Red Balloon* (Kurby et al., 2014; Mumma & Durso, in revision), in which participants segmented the film either into situations or fine events. Furthermore, we also compared the segmentation rate of participants oriented to situations and those oriented to fine events in the present study. Prior to both analyses, the film was divided into 394 5-second time bins. For each participant, each time bin was coded with either a one or a zero, indicating whether a participant reported at least one event boundary within the corresponding 5-second long window of time.

2.2.1.1 Segmentation Agreement

In these analyses, we examined the extent to which the boundaries identified by participants in the present study were related to the normative segmentation of a separate group of participants with the same orientation (i.e., to situations or fine events). In the event segmentation literature, this measure is referred to as "segmentation agreement" and is used widely (e.g., Bailey et al., 2013; Kurby & Zacks, 2011; Newberry & Bailey, 2019; Sargent et al., 2013; Swallow et al., 2018) and is also predictive of psychological outcomes (e.g., Bailey et al., 2013; Kurby & Zacks, 2011).

Using Lorch and Myers' (1990) recommended regression approach for analyzing repeated measures data, we estimated a logistic regression equation separately for each participant that predicted their binary segmentation data simultaneously from the normative fine and situation segmentation data from a separate group of participants. Using the segmentation data from previous studies (Kurby et al., 2014; Mumma & Durso, in revision), a separate normative situation and normative fine segmentation profile was created by dividing the film into 394 5-second time bins and then computing the probability of a person identifying at least one event boundary in each time bin. Both the normative situation and fine segmentation profiles were standardized (i.e., z-score transformed) prior to regression analysis. To assess the reliability of the two logistic regression coefficients over participants from each level of Orientation, we tested whether the mean logistic regression coefficient of each predictor was different from zero using a two-tailed, one-sample *t*-test (with Bonferroni's correction for multiple comparisons).

In Experiment 1, the likelihood of reporting a situation boundary was positively related to the normative probability of identifying a situation boundary (Mean $\beta = 0.65$, $SE = 0.05$), $t(29) = 14.30$, $p < .001$, $d = 1.95$, but not the normative probability of identifying a fine boundary ($M = -0.01$, $SE = 0.05$), $t(29) = -0.17$, $p = 1$, $d = -0.03$. The likelihood of reporting a fine boundary was positively related to the normative probability of identifying a fine boundary ($M = 0.24$, $SE = 0.06$), $t(29) = 4.15$, $p = .002$, $d = 0.76$, and also a situation boundary ($M = 0.48$, $SE = 0.04$), $t(29) = 10.71$, $p < .001$, $d = 1.95$.

2.2.1.2 Segmentation Rate

We expected that situations would be segmented less frequently than fine events because situations are more extended in time (i.e., coarser-grained) than are fine events. This hypothesis is supported from previous research using *The Red Balloon* (Kurby et al., 2014; Mumma & Durso, in revision), which found that participants identified significantly fewer situation boundaries ($M = 74.68$ boundaries, $SE = 5.83$) than fine boundaries ($M = 112.15$ boundaries, $SE = 9.75$) in the film.

For each participant, segmentation rate in each of the four clips was calculated by summing the number of time bins in which a participant identified at least one boundary in a clip and then dividing by the total number of time bins in that clip. We included Clip (1, 2, 3, 4) as a factor to determine whether the segmentation rate of participants with a situation-level orientation became fine-grained over the course of the film.

Segmentation rate was analyzed with a 4 x 2 mixed ANOVA with Clip (1, 2, 3, 4) as a within-subjects factor and Orientation (Situation, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Clip, $F(2.31, 133.93) = 5.55$, $p = .003$, $\eta^2_{partial} = .09$, but no significant effect of Orientation, $F(1, 58) = 3.02$, $p = .088$, $\eta^2_{partial} = .05$, nor interaction between Clip and Orientation, $F(2.31, 133.93) = 1.18$, $p = .32$, $\eta^2_{partial} = .02$ (Figure 3). Because differences between clips were not theoretically interesting, they were not explored further.

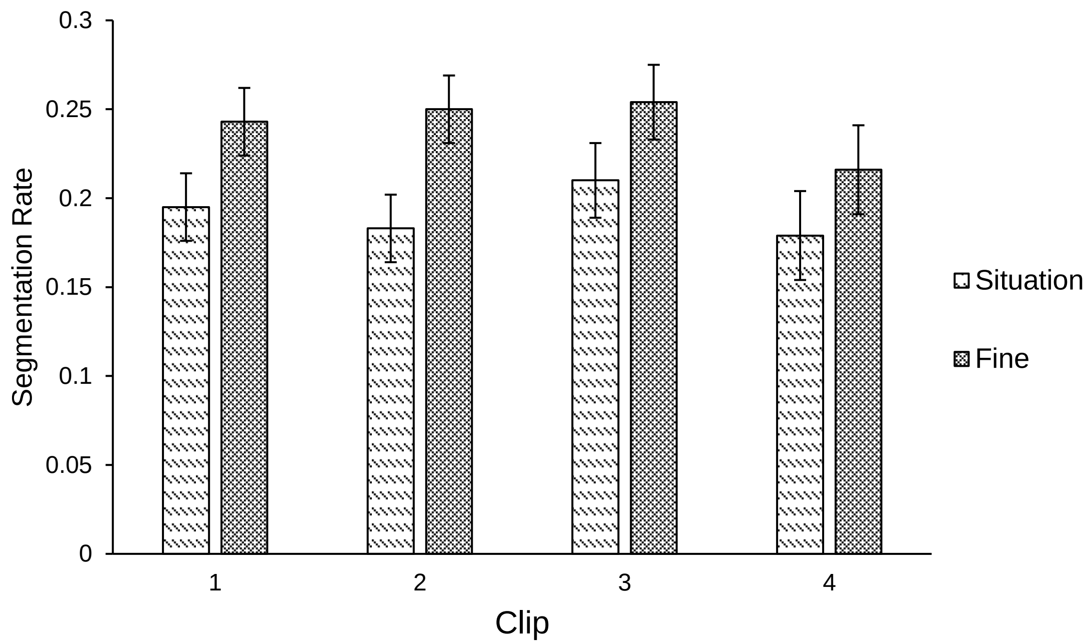

Figure 3 – Mean segmentation rate as a function of Clip and Orientation. Error bars are the standard error of the mean.

In summary, we found that situation segmentation in the present study was uniquely related to the situation segmentation of another group of participants. On the other hand, fine segmentation in the present study was related to normative fine segmentation, but also to normative situation segmentation. That fine segmentation was also related to normative situation segmentation was not surprising given that the beginning of a new situation is very likely to also be the beginning of a new fine event. As we expected, situation segmentation was also consistently coarser in grain than fine segmentation throughout the film (Figure 3), although this difference was not statistically significant by conventional standards (p = .088). Nonetheless, convergence across the two manipulation checks provides more reason to believe that situation segmentation differed from fine segmentation in the expected direction than not.

2.2.2 Neutral-Grain Segmentation

Neutral-grain segmentation reflects the grain at which participants naturally (or preferentially) segmented the film. Like fine segmentation, neutral-grain segmentation was also positively related to both the normative probability of a situation boundary (M = 0.54, SE = 0.07), $t(29) = 7.34$, $p < .001$, $d = 1.34$, and a fine boundary (M = 0.21, SE = 0.05), $t(29) = 4.10$, $p = .002$, $d = 0.75$.

To compare the segmentation rates of participants under each Orientation, segmentation rate was analyzed with a 4 x 3 mixed ANOVA with Clip (1, 2, 3, 4) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Clip, $F(2.13, 185.40) = 9.10$, $p < .001$, $\eta^2_{partial} = .10$, but no significant effect of Orientation, $F(2, 87) = 1.88$, $p = .16$, $\eta^2_{partial} = .04$, nor interaction between Clip and Orientation, $F(4.26, 185.40) = 0.85$, $p = .50$, $\eta^2_{partial} = .02$ (Figure 4). Thus, although the analysis of segmentation agreement suggested that neutral-grain segmentation was more like fine than situation segmentation, analysis of segmentation rates could not clearly distinguish neutral-grain segmentation from either situation or fine segmentation.

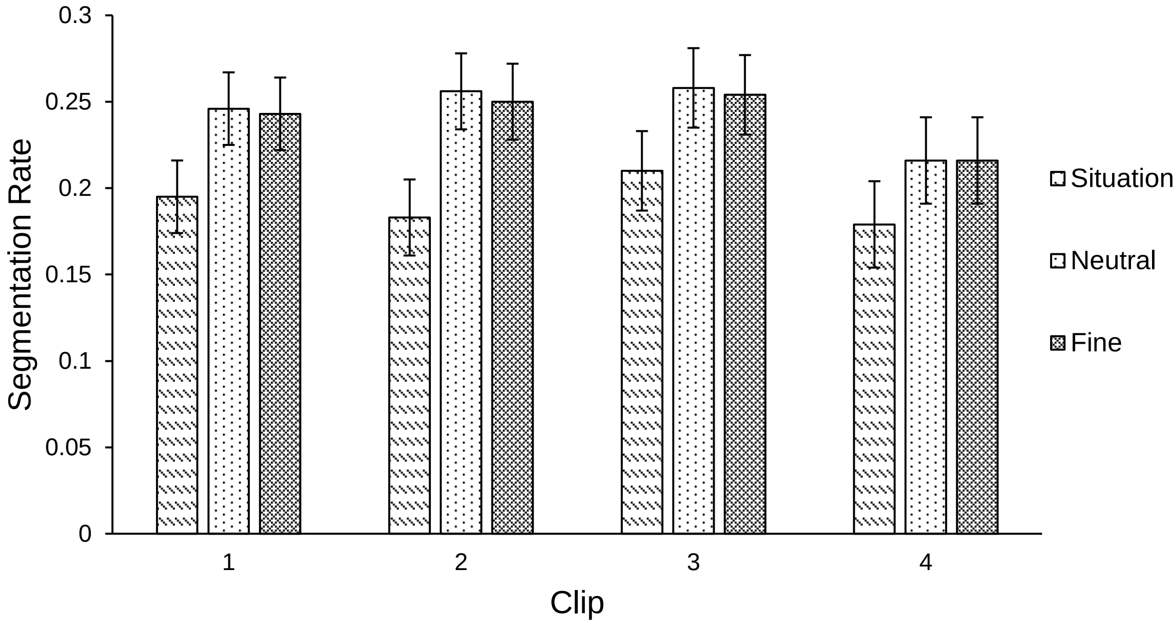

Figure 4 – Mean segmentation rate as a function of Clip and Orientation. Error bars are the standard error of the mean.

2.2.3 Confidence Ratings

Mean confidence ratings were subjected to a 3 x 3 mixed ANOVA with Probe Type (Within/Within, Within/Across, Across/Across) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Probe Type, $F(2, 174) = 34.61$, $p < .001$, $\eta^2_{partial} = .29$, but no significant effect of Orientation, $F(2, 87) = 0.14$, $p = .87$, $\eta^2_{partial} = .003$, nor interaction between Probe Type and Orientation, $F(4, 174) = 1.64$, $p = .17$, $\eta^2_{partial} = .04$.

We examined the main effect of Probe Type using orthogonal polynomial contrasts (linear and quadratic) because the levels of Probe Type correspond to an incrementally increasing hierarchical event structure - zero event boundaries (Within/Within), one event boundary (i.e., a fine event boundary; Within/Across), and two event boundaries (i.e., a fine event plus a situation boundary; Across/Across). Trend analysis of Probe Type indicated a significant linear, $F(1, 87) = 53.05$, $p < .001$, $\eta^2_{partial} = .38$, as well as quadratic trend, $F(1, 87) = 7.36$, $p = .008$, $\eta^2_{partial} = .08$. These trends suggest that confidence in one's ability to predict what happens next is greatest when within a situation and fine event (Within/Within; $M = 3.45$, $SE = 0.09$), but is similarly low when nearing either the end of a fine event in an ongoing situation (Within/Across; $M = 3.09$, $SE = 0.08$) or the end of both a fine event and a situation (Across/Across; $M = 2.97$, $SE = 0.08$; Figure 5).

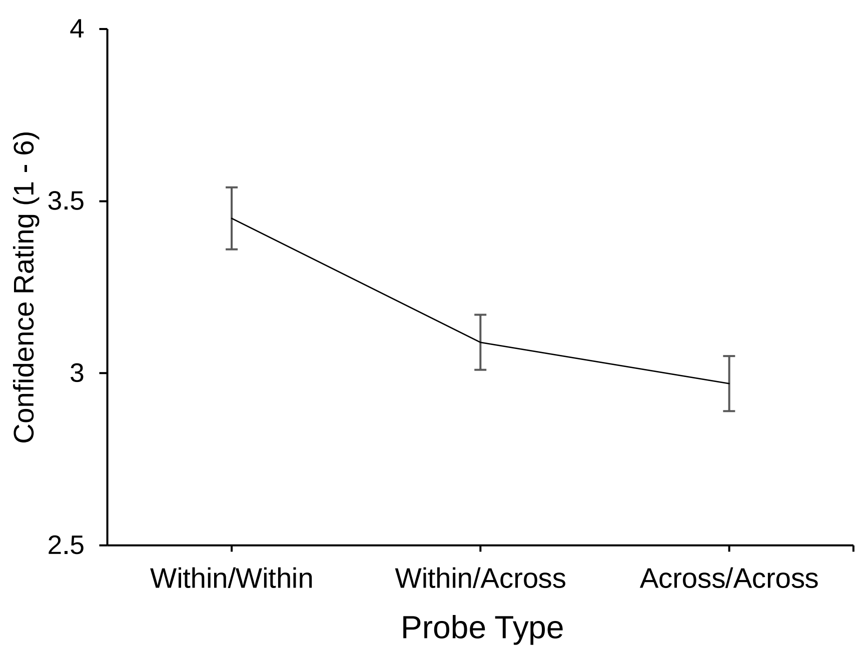

Figure 5 – Mean confidence rating as a function of Probe Type. Error bars are the standard error of the mean.

2.2.3.1 Confidence and Segmentation

Event segmentation theory proposes that an error-detection monitoring system regulates event model updating and that prediction error is associated with subjective uncertainty (Richmond & Zacks, 2017; Zacks et al., 2011). Consequently, confidence in one's ability to predict what happens next should be negatively related to the likelihood of perceiving an event boundary in the near future (i.e., shortly after the resumption of the film after a probe). To determine whether a participant segmented after each probe, we established a lenient response window (as probes may have been locally disruptive to event segmentation). Based on the mean response latency in the 10 seconds following boundary-probes (i.e., Within/Across and Across/Across probes), we included only key-presses that fell within 5.29 seconds ± 1.96 standard deviations (*SD* = 2.41) after each probe location. However, we excluded key-presses falling within 1.96 standard deviations before the next closest consensual boundary at the same level of Orientation (e.g., the next consensual situation boundary if segmenting the film into situations); for neutral-grain segmentation, we used the next closest consensual fine boundary.

The relationship between confidence ratings (standardized separately for each participant) and the report of event boundaries shortly after each probe location was examined using the Lorch and Myers (1990) procedure described previously. Briefly, for each participant, we estimated a logistic regression equation that predicted that participant's binary segmentation data from his or her confidence ratings. Given the small number of observations for each participant, logistic regression coefficients were estimated using penalized maximum likelihood rather than conventional maximum likelihood, which may produce biased estimates in small samples (Allison, 1999).

Logistic regression coefficients for 5 participants (17%) in the Neutral orientation and 2 participants (7%) in the Fine orientation could not be estimated because these participants did not segment after any of the probes. Using the remaining participants, a one-way ANOVA with Orientation as a between-subjects factor revealed no differences among logistic regression coefficients between levels of Orientation, $F(2, 80) = 0.21$, $p = .81$, $\eta^2_{partial} = .01$. Consequently, data were collapsed across levels of Orientation and collectively compared to zero using a one-sample *t*-test. The mean logistic regression coefficient ($M = -0.24$, $SE = 0.07$) was significantly less than zero, $t(82) = -3.54$, $p < .001$, $d = -0.39$. Thus, the more confident one is in his or her ability to predict what will happen next, the less likely he or she is to report an event boundary in the near future.

2.2.4 Consistency Ratings

The reliability of the average consistency rating for each level of Probe Type was assessed with three separate intraclass correlation coefficients (ICCs), which used an average measures absolute agreement definition (ICC [2, k]; Shrout & Fleiss, 1979). According to conventional guidelines for interpreting the values of ICCs (Cicchetti, 1994), the ICCs ranged from good (Within/Within ICC = .62; 95% CI = .43 - .75) to excellent (Within/Across ICC = .78; 95% CI = .69 - .85 and Across/Across ICC = .82; 95% CI = .75 - .88).

Mean consistency ratings were subjected to a 3 x 3 mixed ANOVA with Probe Type (Within/Within, Within/Across, Across/Across) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Probe Type, $F(1.64, 143.03) = 37.58$, $p < .001$, $\eta^2_{partial} = .30$, but no significant effect of Orientation, $F(2, 87) = 1.53$, $p = .22$, $\eta^2_{partial} = .03$, nor interaction between Probe Type and Orientation, $F(3.29, 143.03) = 0.28$, $p = .89$, $\eta^2_{partial} = .01$.

Trend analysis of Probe Type indicated a significant linear, $F(1, 87) = 48.47$, $p < .001$, $\eta^2_{partial} = .36$, as well as quadratic trend, $F(1, 87) = 24.66$, $p < .001$, $\eta^2_{partial} = .22$. Thus, predictions made within an ongoing situation and fine event (Within/Within; $M = 2.95$, $SE = 0.03$) are most consistent with what happens in the very near future compared to predictions made when nearing either the end of a fine event in an ongoing situation (Within/Across; $M = 2.60$, $SE = 0.03$) or the end of both a fine event and situation (Across/Across; $M = 2.62$, $SE = 0.05$; Figure 6).

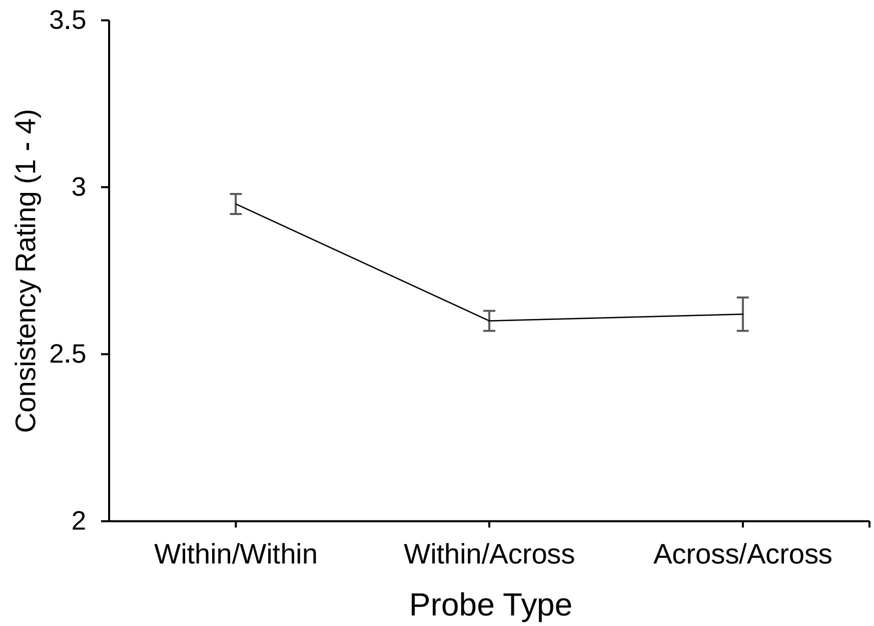

Figure 6 – Mean consistency rating as a function of Probe Type. Error bars are the standard error of the mean.

<u>2.2.4.1 Consistency and Confidence</u>

To examine the relationship between confidence and consistency ratings, a linear regression equation that predicted consistency ratings from confidence ratings (standardized for each participant separately) was computed for each participant (Lorch & Myers, 1990). A one-way ANOVA with Orientation as a between-subjects factor revealed no differences in regression coefficients between levels of Orientation, $F(2, 87) = 0.43$, $p = .65$, $\eta^2_{partial} = .01$. Consequently, data were collapsed across levels of Orientation and collectively compared to zero using a one-sample *t*-test. The mean regression coefficient was small but statistically greater than zero ($M = 0.10$, $SE = 0.02$), $t(89) = 4.51$, $p < .001$, $d = 0.48$. Thus, the more confidently one can predict what happens next, the more consistent one's prediction is with what actually happens in the near future.

2.2.4.2 Consistency and Segmentation

Given that confidence ratings were negatively related to the likelihood of reporting an event boundary in the near future, we determined whether the accuracy of one's prediction was also negatively related to the likelihood of reporting an event boundary in the near future. As before, a separate logistic regression equation was estimated for each subject using penalized maximum likelihood estimation. Logistic regression coefficients for the same 5 participants in the Neutral orientation and 2 participants in the Fine orientation could not be estimated because these participants did not report a boundary after any probe.

Using the remaining participants, a one-way ANOVA with Orientation as a between-subjects factor revealed no difference in the logistic regression coefficients among levels of Orientation, $F(2, 80) = 1.64$, $p = .20$, $\eta^2_{partial} = .04$. Consequently, data were collapsed across levels of Orientation and collectively compared to zero using a one-sample *t*-test. The mean logistic regression coefficient ($M = -0.25$, $SE = 0.07$) was significantly less than zero, $t(82) = -3.52$, $p < .001$, $d = -0.39$. Thus, the more consistent one's prediction is with the near future, the less likely he or she is to report an event boundary in the near future.

2.3 Discussion

In summary, Experiment 1 provided converging evidence for the general mechanisms of EST, but evidence against the hypothesis that the hierarchical structure of event segmentation is attributable to the differential predictive accuracies of event models on different timescales (i.e., of situations versus fine events).

Regarding the former, EST proposes that event models are updated in response to transient increases in prediction error, which is associated with one's subjective uncertainty about the future (Zacks et al., 2011). It follows that confidence in one's ability to predict the future and the accuracy of one's prediction should each be negatively related to the likelihood of perceiving an event boundary (i.e., updating the current event model) in the near future. Indeed, we found that the more confident about and the more consistent one's prediction is with what happens in the near future, the less likely one is to subsequently report an event boundary. Moreover, the more confident one is in their ability to predict the future, the more consistent one's prediction is with what happens next.

However, Experiment 1 failed to provide evidence for the differential predictive accuracies of event models on different timescales. Regardless of how participants were oriented to activity in the film, both their confidence in their ability to predict the future and the extent to which their predictions were consistent with what happened in the near future followed a pattern that we expected (at least) for fine segmentation (see Figure 1). Specifically, confidence and consistency were highest when assessed within a fine event in an ongoing situation (Within/Within) but were equally low when assessed either near the end of a fine event in an ongoing situation (Within/Across) or near the end of a fine event and the current situation (Across/Across). This pattern suggests that, regardless of one's orientation to activity, confidence and predictive accuracy vary as a function of whether a fine event boundary is imminent or not.

The failure to observe an effect of Orientation on confidence or predictive accuracy cannot be attributed to a failure to manipulate Orientation. Importantly, situation and fine segmentation were differentially related to the normative situation and fine segmentation of a different group of participants who segmented *The Red Balloon* into either situations or fine events. Consistent with our expectations, situations were also segmented more coarsely than were fine events throughout the film, although statistical support for this observation was not as strong. Otherwise, neutral-grain segmentation bared some resemblance to fine segmentation, suggesting that participants may have preferentially segmented the film into small events.

In short, the results of Experiment 1 suggest that participants had event models of activity only at the fine-grained level. In support, both confidence and predictive accuracy varied simply as a function of whether a fine event boundary was imminent or not, regardless of one's orientation. Moreover, this pattern was observed even though participants oriented to situations segmented the film differently than participants oriented to either neutral-grain or fine events.

CHAPTER 3. EXPERIMENT 2

Methodologically, Experiment 2 was identical to Experiment 1 except that we assessed predictive accuracy (i.e., discriminability) using a Yes/No decision task rather than an explicit prediction task. Using a different operationalization of predictive accuracy allowed us to test the generalizability of our findings from Experiment 1, consistent with the three-pronged method for studying inference generation (Magliano & Graesser, 1991) and also with how previous studies have assessed the effect of event structure on predictive accuracy (Huff et al., 2014; Zacks et al., 2011).

Unlike previous studies (Huff et al., 2014; Zacks et al., 2011), however, we used brief verbal statements as targets and foils rather than pictures (i.e., video-stills). Doing so circumvented the challenges of creating pictorial targets and foils from a narrative film (e.g., equating the visual or cinematic features of targets and foils) and moreover, a verbal statement can convey extended activity less ambiguously than a single static picture. Lastly, we did not assess the relationship between confidence ratings and accuracy and the likelihood of segmentation and accuracy (as we did in Experiment 1) because it is not possible to separate the contribution(s) of discriminability and bias to accuracy (i.e., correct or incorrect) at the trial (or probe) level (Gronlund & Neuschatz, 2014).

3.1 Method

3.1.1 Participants

We targeted a sample size of 90 participants (30 in each level of Orientation) to detect a moderately sized interaction (Cohen's f = .25) between Probe Type and Orientation with 80% power. Eighty-eight undergraduate students from the Georgia Institute of Technology participated in partial fulfillment of a research familiarization requirement. All data for four participants were excluded because two participants reported that they had seen the film before on the post-experiment questionnaire and two participants failed to follow instructions consistently during the experiment. These participants were replaced. The remaining 84 participants had a mean age of 20 years (*SD* = 1.5) and were 52% male.

3.1.2 Design

Like Experiment 1, Experiment 2 used a 3 (Orientation: Situation, Neutral, Fine) x 3 (Probe Type: Within/Within, Within/Across, Across/Across) mixed design with Orientation varying between subjects and Probe Type varying within subjects. For each level of Probe Type, we measured each participant's confidence in his or her ability to predict what would happen next in the film, how quickly and accurately he or she could identify what would actually happen next, as well as his or her response bias. We also recorded the exact moment in time that participants reported an event boundary during the film.

3.1.3 Materials

3.1.3.1 Targets

The target for each probe was a brief description of what happened during the 5 seconds of film immediately following the probe. To ensure that the descriptions of targets would not be idiosyncratic, three judges independently watched and described (in 5 – 10 words) what happened in the 5 seconds following each probe location. For each probe location, a fourth judge (JM) produced a single description that captured the common (i.e., consensual) elements of the three descriptions. At one probe location, for example, the three judges described the corresponding clip as, "Boy climbs down post, grabs balloon and bag, then walks away," "Boy picks up bag, holds balloon in mouth, keeps walking," and "Boy walks away from lamp post with the balloon." Accordingly, the fourth judge summarized these three descriptions as "The boy walks away with the balloon."

3.1.3.2 Foils

From the 1,620 predictions generated in Experiment 1, our goal was to create a foil for each probe location, which would be based on an incorrect prediction that was infrequently generated under each level of Orientation. We opted to use infrequent (rather than frequent) predictions to avoid obtaining a floor effect in predictive accuracy. To this end, we used a card sorting procedure to reduce the 1,620 predictions into a smaller number of "universal predictions" at each probe location (i.e., similar predictions that were generated under each level of Orientation).

For each probe location, three judges independently sorted the 30 predictions into piles such that individual predictions within the same pile were more similar to each other than individual predictions from different piles. Judges were allowed to make as many or as few piles as they felt was necessary, with each pile containing at least one prediction. For each pile containing more than one prediction, judges created a short descriptive label that captured what the constituent predictions had in common (i.e., formed categories of predictions). At one probe location, for example, one judge placed the predictions “He buys a pastry,” “The boy buys some dessert,” and “The boy will buy a treat for himself,” together in a pile labelled “The boy buys some food.” The judges repeated this entire process three times, once for each level of Orientation.

To identify consensual categories at each probe location, a fourth judge (JM), who did not participate in the original card sort, examined the category labels given to each prediction by the three judges. If at least two of the three judges categorized a prediction under a similar label, the fourth judge created a new category label that captured what the category labels had in common. At one probe location, for example, the three original judges placed the prediction “The boy will keep walking to wherever he is going,” in a pile of predictions labeled “Boy keeps walking,” “The kid keeps walking,” and “Boy continues walking”, respectively. Thus, for that prediction, the fourth judge created a consensual label entitled, “The boy continues walking.” For each probe location under each level of Orientation, the fourth judge generated a list of consensual category labels with each label applying to at least two predictions at a probe location. The median number of consensual categories identified per probe location was 5 (Range: 2 – 8), 5 (Range: 3 – 7), and 4.5 (Range: 3 – 7) for Situation, Neutral, and Fine orientations, respectively.

Next, all four judges independently sorted the predictions at each probe location using only the consensual category labels for that probe location identified by the fourth judge. Judges labelled any prediction that did not fit into a category label as “Other.” The judges repeated this process for each level of Orientation. Agreement was achieved if at least three of the four judges placed a given prediction in the same category. On average, judges achieved consensus for 98% (SD = 0.02), 97% (SD = 0.04), and 95% (SD = 0.05) of predictions at each probe location under Situation, Neutral, and Fine orientations, respectively. The few predictions on which the judges disagreed were each resolved via discussion among the group. For each level of Orientation, the card sorting process dramatically reduced the 540 predictions across all probe locations to 184, 210, and 187 unique predictions in the Situation, Neutral, and Fine orientations, respectively (see Appendix A for a table of predictions for each probe location).

Using the reduced list of predictions, we proceeded to identify predictions at each of the 18 probe locations that appeared under all three levels of Orientation (i.e., “universal predictions”). To this end, a single judge (JM) identified similar predictions that appeared under each level of Orientation (i.e., “constituent predictions”) and then labelled the universal prediction accordingly. Universal predictions could be made of categories of predictions, lone predictions that did not fit into any category under a certain level of Orientation (i.e., a singleton), or both. At one probe location, for example, the universal prediction “The balloon will get the key to the door,” comprised a category of predictions from two levels of Orientation (“The balloon will get the key to the door,” and “The balloon will get the key to the door.”) and one singleton from the other level of

Orientation ("The balloon will try to get the key and rescue the boy."). Agreement across the 18 probe locations was very high, with the majority of the judges identifying the same three constituent predictions for 94% (62/66) of the universal predictions. As before, the few universal predictions that failed to reach consensus were each resolved via discussion among the group. The median number of universal predictions at each probe location was 3.5 (Range: 2 – 6).

To create a foil for each of the 18 probe locations, we drew from each probe location's respective set of universal predictions. Ultimately, our goal was to select universal predictions for each probe location that 1) did not come true in the film, 2) were low in consistency with what happens shortly after a probe (i.e., less than 3, which corresponds to "Moderately consistent"), and 3) that only a small and similar number of participants made under each level of Orientation (i.e., less than 25% in each level of Orientation).

For each of the three constituent predictions of a universal prediction, we obtained the percentage of participants who made that prediction. At one probe location, for example, the universal prediction "The children try to grab the balloon," comprised "Children will try to catch the balloon," with 7% of participants with a situation-level orientation making this prediction, "Children will try to catch the balloon," which was made by 7% of those with a neutral orientation, and "The kids try and grab the balloon," which was made by 3% of those with a fine orientation. To obtain the consistency rating of a constituent prediction, we first obtained the consistency rating for each individual prediction (averaged over raters) in a constituent prediction and then averaged those values together. Given the criteria above, we ultimately selected universal predictions with constituent predictions that did not come true and were similarly infrequent and low in their consistency with what happened next in the film. Table 4 shows the percentage of all participants making each universal prediction (i.e., foil) as well as the average consistency rating of each universal prediction.

Lastly, the number of syllables for each target-foil pair were matched as closely as possible. Across the 18 probe locations, the median number of syllables for targets and foils was 9 with a range of 6 – 11 and 5 – 11, respectively. For all participants, half of the 18 probes were targets (i.e., the correct answer was "Yes"). In each level of Orientation, however, we randomly assigned half of the participants to one random sequence of targets and foils and the other participants to a sequence with the target and foil-assignments reversed.

Table 4 – Targets and foils for each probe location, with the percentage of all participants making each universal prediction (i.e., foil) and the mean consistency of the foil.

Probe	Probe Type	Target	Foil	Percentage	Consistency
1	Within/Across	The boy walks towards the stairs.	The cat runs away.	10%	2.22
2	Within/Across	The boy walks away with the balloon.	The boy loses the balloon.	7%	1.89
3	Across/Across	The boy runs to a door.	The boy gets injured.	3%	1.44
4	Within/Within	The boy and the man walk together.	The boy and the man go into a store.	8%	2.62
5	Across/Across	The boy releases the balloon outside.	The boy ties the balloon to the balcony.	3%	1.33
6	Within/Within	The boy and the balloon wait to board the bus.	The boy and the balloon board the bus. [a]	2%	2.34
7	Within/Across	The boy gets off the bus.	The balloon has trouble following the bus.	6%	1.93
8	Across/Across	The balloon escapes from the man.	The man catches the balloon.	19%	2.14
9	Within/Within	The balloon follows the man.	The children try to grab the balloon.	6%	2.27
10	Within/Across	The men shake hands and part ways.	The balloon takes the key from the man.	4%	2.25
11	Across/Across	The boy and the balloon reunite.	A stranger steals the balloon.	3%	1.22
12	Within/Across	The boy and the girl part ways.	The boy and the girl play together.	9%	1.71
13	Across/Across	The boy walks with his mother.	The boy's mother puts the balloon outside.	11%	1.60

Table 4 (continued).

14	Within/Within	The boy checks his pockets for money.	The boy uses the balloon to get food.	6%	2.47
15	Within/Across	The boy and the balloon hide in a doorway.	The girl with the blue balloon appears. [b]	7%	2.31
16	Across/Across	A boy fatally stomps on the balloon.	The girl with the blue balloon appears. [b]	6%	2.40
17	Within/Within	The balloons fly over the city.	The balloons retaliate against the boys.	12%	2.61
18	Within/Within	The balloons lift the boy into the sky.	The boys see the boy flying with the balloons.	11%	2.57

[a] This was the only foil based on a universal prediction that was generated under two of the three levels of Orientation.

[b] Although this foil was used twice, participants only encountered this foil once in the two randomized sequences of targets and foils.

3.1.3.3 Performance Measures

For Yes/No tasks, there are a number of ways in which performance can be measured. We avoided performance measures that confound discriminability and response bias (e.g., percentage correct; Stanislaw & Todorov, 1999) in favor of indices that measure discriminability and bias directly and independently of each other. For Yes/No tasks, common measures of discriminability include *d'*, log *d* (Brown & White, 2005; or log α, Macmillan & Creelman, 1990), and A' (Pollack & Norman, 1964). Despite its widespread usage, A' is generally not recommended as a measure of discriminability (Macmillan & Creelman, 1996; Macmillan, Rotello, & Miller, 2004;

Pastore, Crawley, Berens, & Skelly, 2003; Rhodes, Cowan, Parra, & Logie, 2018; Rotello, Masson, & Verde, 2008; Stanislaw & Todorov, 1999; Verde, Macmillan, & Rotello, 2006). Between *d'* and log *d*, we chose the latter index because *d'* is not recommended when the number of signal and noise trials are both small (i.e., N < 100 each; Brown & White, 2005; Kadlec, 1999), whereas Brown and White (2005) recommend log *d* in such cases. Like *d'*, log *d* (Equation 1; Macmillan & Creelman, 1990) can theoretically range in value from 0, corresponding to the complete inability to discriminate between signal and noise (i.e., chance performance), to positive infinity. We chose a complementary measure of response bias (i.e., criterion location), log *b* (Equation 2; Macmillan & Creelman, 1990), that is independent of log *d* (Corwin, 1994; Macmillan & Creelman, 1990; Snodgrass & Corwin, 1988). For log *b*, negative values correspond to a liberal response bias (i.e., a greater tendency to say "Yes" than "No") and positive values correspond to a conservative response bias (i.e., a greater tendency to say "No" than "Yes").

$$\text{Log } d = \frac{1}{2} \cdot \{\log_{10}[\frac{Hits}{(1-Hits)}] - \log_{10}[\frac{False\ Alarms}{(1-\ False\ Alarms)}]\} \tag{1}$$

$$\text{Log } b = -\frac{1}{2} \cdot \{\log_{10}[\frac{Hits}{(1-Hits)}] + \log_{10}[\frac{False\ Alarms}{(1-\ False\ Alarms)}]\} \tag{2}$$

Lastly, log *d* is indeterminate or infinite when a pair of hit and false alarm rates contains an extreme value (i.e., either 1 or 0). Extreme pairs (i.e., pairs with at least one extreme value) were observed for 69% of all participants for Within/Within probes, 46% for Within/Across probes, and 36% for Across/Across probes. The frequency of extreme pairs did not differ significantly between levels of Orientation for Within/Within, $\chi^2(2)$ = 1.78, p = .41, Within/Across, $\chi^2(2)$ = 5.46, p = .07, nor Across/Across, $\chi^2(2)$ = 2.18, p = .34. A loglinear transformation was applied to all hit and false alarm rates, regardless if they were extreme in value (Brown & White, 2005; Hautus, 1995; Kadlec, 1999; Snodgrass & Corwin, 1988; Stanislaw & Todorov, 1999). This transformation involves adding 0.5 to both the number of hits and false alarms and 1 to the number of signal and noise trials (Equation 3).

$$\text{Loglinear Transformation} = \frac{\#Hits + 0.5}{\#Signal\ Trials + 1}, \frac{\#False\ Alarms + 0.5}{\#Noise\ Trials + 1} \tag{3}$$

3.1.4 Procedure

The procedure for Experiment 2 was nearly identical to those in Experiment 1, but with two exceptions. First, participants were not only randomly assigned to a level of Orientation, but also to one of two random sequences of targets and foils. Second, rather than writing down a prediction of what happens next in the film at each probe location, participants were given one of two options (i.e., either a target or a foil) and asked, “Does the statement below describe what will happen next?” Participants were instructed not to overthink their answers and to respond as quickly, but also as accurately, as possible. The experimental program recorded each participant’s confidence ratings, their response and response latency to probes (i.e., yes or no), and the exact moment(s) he or she pressed the spacebar in between probes.

3.1.5 Analyses

For all analyses, the criterion for significance was an α level of .05. When corrections for multiple comparisons were used (e.g., the Bonferroni correction), the reported *p* values are adjusted and can be compared to .05. The violation of sphericity in mixed model ANOVAs was assessed with Mauchly's sphericity test and if violated, degrees of freedom were adjusted using the Greenhouse-Geisser correction. As estimates of effect size, we report partial eta-squared ($\eta^2_{partial}$) for ANOVA effects and Cohen's *d* for *t*-tests.

3.2 Results

3.2.1 Manipulation Check

As in Experiment 1, we examined the degree to which situation and fine segmentation in Experiment 2 were related to normative segmentation data from previous segmentation studies using *The Red Balloon* (Kurby et al., 2014; Mumma & Durso, in revision). Additionally, we compared the segmentation rate of participants oriented to situations and those oriented to fine events in the present study. Prior to both analyses, the film was divided into 394 5-second time bins. Each time bin was coded for whether a participant reported at least one event boundary within the corresponding 5-second window of time.

3.2.1.1 Segmentation Agreement

Using the same analytical approach as in Experiment 1 (Lorch & Myers, 1990), we examined the extent to which the boundaries identified by participants in Experiment 2 were related to the normative segmentation of another group of participants with the same orientation. The likelihood of reporting a situation boundary in Experiment 2 was positively related to the normative probability of identifying a situation boundary (Mean $\beta = 0.59$, $SE = 0.06$), $t(27) = 10.07$, $p < .001$, $d = 1.95$, but not the normative probability of identifying a fine boundary ($M = 0.09$, $SE = 0.06$), $t(27) = 1.34$, $p = 1$, $d = 0.25$. The likelihood of reporting a fine boundary was positively related to both the normative probability of identifying a fine boundary ($M = 0.26$, $SE = 0.07$), $t(27) = 3.93$, $p = .003$, $d = 0.74$, as well as a situation boundary ($M = 0.52$, $SE = 0.05$), $t(27) = 9.75$, $p < .001$, $d = 1.84$.

3.2.1.2 Segmentation Rate

Segmentation rate was analyzed with a 4 x 2 mixed ANOVA with Clip (1, 2, 3, 4) as a within-subjects factor and Orientation (Situation, Fine) as a between-subjects factor. As in Experiment 1, including Clip as a factor allowed us to determine whether the segmentation rate of participants with a situation-level orientation became fine-grained over time. This analysis revealed a significant main effect of Clip, $F(2.20, 118.84) = 10.52$, $p < .001$, $\eta^2_{partial} = .16$, a significant main effect of Orientation, $F(1, 54) = 12.95$, $p < .001$, $\eta^2_{partial} = .19$, and no significant interaction between Clip and Orientation, $F(2.20, 118.84) = 1.82$, $p = .16$, $\eta^2_{partial} = .03$ (Figure 7). Because differences between clips were not theoretically interesting, they were not explored further. Consistent with our expectations, the main effect of Orientation suggests that situations were segmented at a significantly lower rate ($M = 0.20$, $SE = 0.03$) than were fine events ($M = 0.34$, $SE = 0.03$) in Experiment 2.

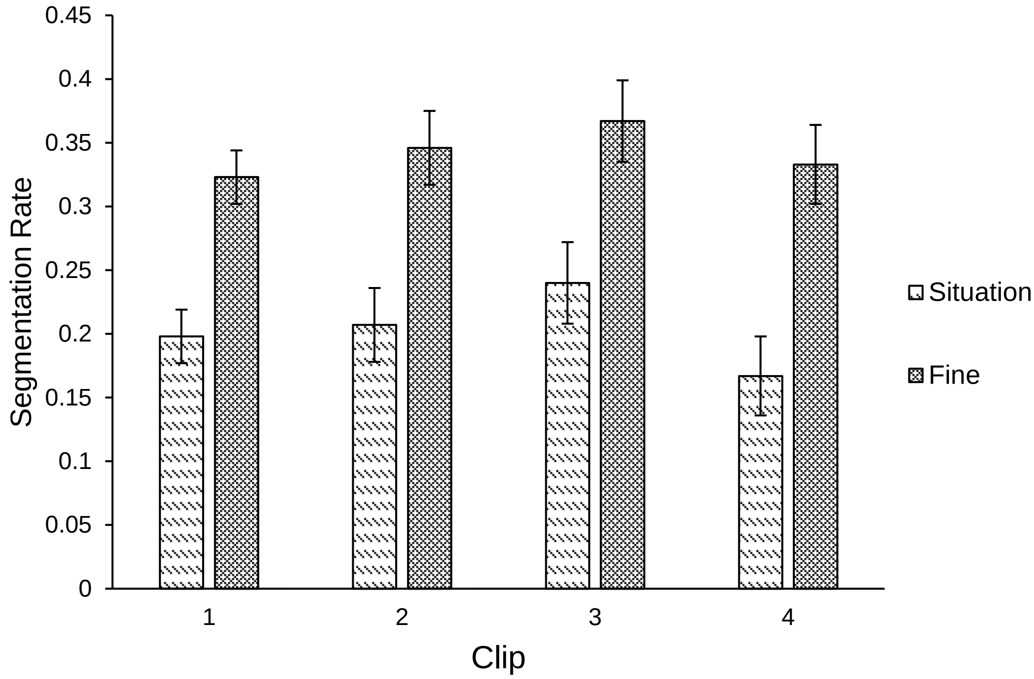

Figure 7 – Mean segmentation rate as a function of Clip and Orientation. Error bars are the standard error of the mean.

3.2.2 Neutral-Grain Segmentation

Like situation segmentation, neutral-grain segmentation in Experiment 2 was positively related only to the normative probability of identifying a situation boundary (*M* = 0.61, *SE* = 0.05), $t(27) = 11.42$, $p < .001$, $d = 2.16$, and not the normative probability of identifying a fine boundary (*M* = -0.01, *SE* = 0.07), $t(27) = -0.11$, $p = 1$, $d = -0.02$. To compare the segmentation rates of participants under each Orientation, we analyzed segmentation rate with a 4 x 3 mixed ANOVA with Clip (1, 2, 3, 4) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Clip, $F(2.34, 189.30) = 12.25$, $p < .001$, $\eta^2_{partial} = .13$, a significant main effect of Orientation, $F(2, 81) = 11.37$, $p < .001$, $\eta^2_{partial} = .22$, and a significant interaction between Clip and Orientation, $F(4.67, 189.3) = 2.60$, $p = .03$, $\eta^2_{partial} = .06$.

Because our concern was whether differences in segmentation rate between the levels of Orientation changed over time, analysis of the interaction between Clip and Orientation focused on the simple main effect of Orientation for each of the four clips (using Bonferroni's correction for multiple comparisons). These analyses suggested that in each of the four clips, participants with a fine orientation always segmented at a significantly higher rate than did participants with either a neutral-grain or a situation orientation (largest $p = .04$). Otherwise, the segmentation rate for participants with either a neutral-grain or a situation orientation did not differ for any clip (all ps = 1). Thus, differences in the segmentation rate of participants under different levels of Orientation remained constant throughout the film (Figure 8).

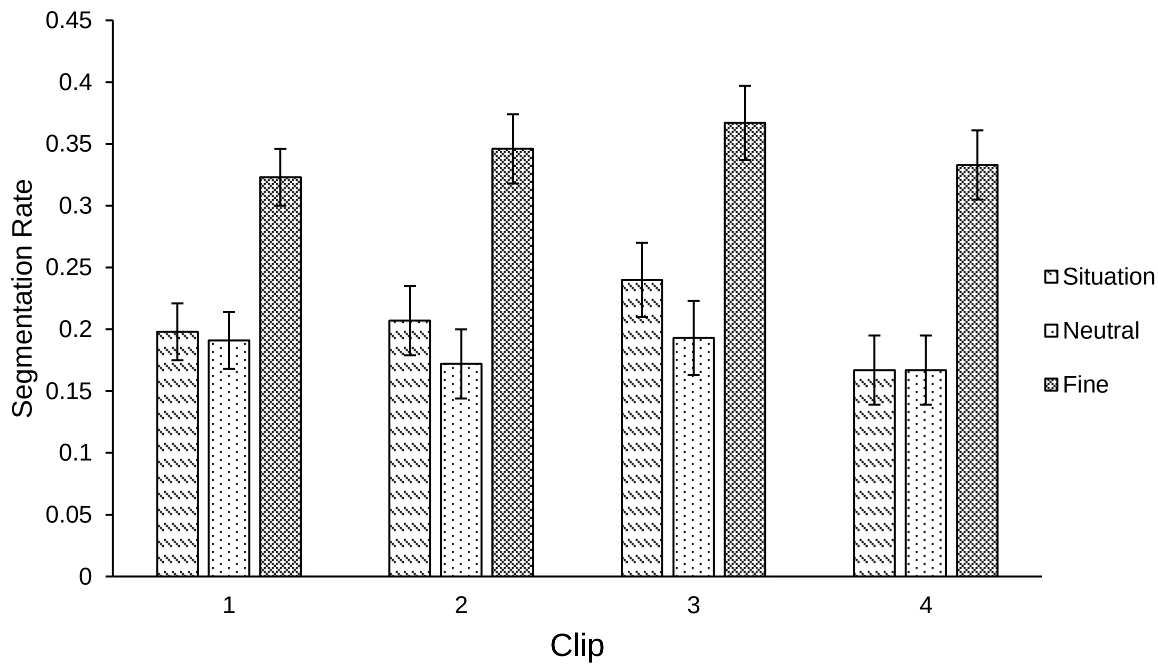

Figure 8 – Mean segmentation rate as a function of Clip and Orientation. Error bars are the standard error of the mean.

3.2.3 Confidence Ratings

Mean confidence ratings were subjected to a 3 x 3 mixed ANOVA with Probe Type (Within/Within, Within/Across, Across/Across) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Probe Type, $F(2, 162) = 36.29$, $p < .001$, $\eta^2_{partial} = .31$, but no significant effect of Orientation, $F(2, 81) = 0.19$, $p = .83$, $\eta^2_{partial} = .01$, nor interaction between Probe Type and Orientation, $F(4, 162) = 0.38$, $p = .83$, $\eta^2_{partial} = .01$.

A trend analysis of Probe Type indicated a significant linear, $F(1, 81) = 55.84$, $p < .001$, $\eta^2_{partial} = .41$, as well as a quadratic trend, $F(1, 81) = 11.33$, $p = .001$, $\eta^2_{\mathrm{partial}} = .12$.

Thus, confidence in one's ability to predict what happens next is greatest when within an ongoing situation and fine event (Within/Within; *M* = 3.70, *SE* = 0.09), but is similarly low when nearing either the end of a fine event in a situation (Within/Across; *M* = 3.34, *SE* = 0.09) or the end of a fine event and situation (Across/Across; *M* = 3.27, *SE* = 0.09; Figure 9).

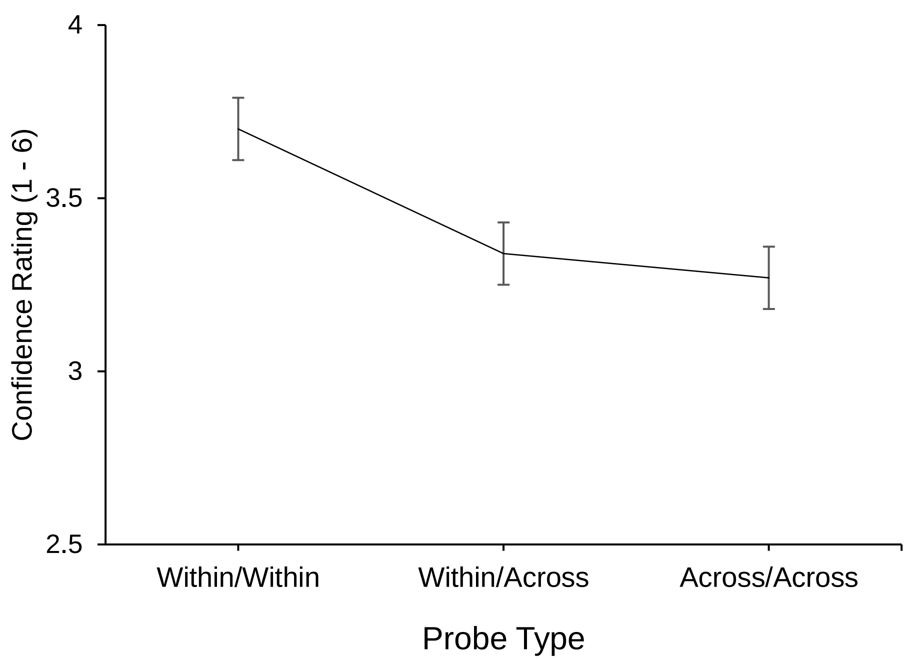

Figure 9 – Mean confidence rating as a function of Probe Type. Error bars are the standard error of the mean.

<u>3.2.3.1 Confidence and Segmentation</u>

To assess the relationship between confidence and the likelihood of segmentation, we repeated the same analysis for Experiment 2 that we followed in Experiment 1.

Briefly, we estimated a logistic regression equation for each participant predicting his or her binary segmentation data following a probe from the corresponding confidence rating (standardized within each participant separately). To determine whether a participant segmented after each probe, we established a lenient response window (as probes may have been locally disruptive to event segmentation). Based on the mean response latency in the 10 seconds following boundary-probes (i.e., Within/Across and Across/Across probes), we included only key-presses that fell within 5.04 seconds ± 1.96 standard deviations (SD = 2.51) after each probe location. However, we excluded key-presses falling within 1.96 standard deviations before the next closest consensual boundary at the same level of Orientation (e.g., the next consensual situation boundary if segmenting the film into situations); for neutral-grain segmentation, we used the next closest consensual fine boundary.

Logistic regression coefficients for 5 participants (18%) with a Neutral orientation, 2 participants (7%) with a Fine orientation, and 1 participant (4%) with a Situation orientation could not be estimated because these participants did not segment after any of the probes. Additionally, data for 1 participant with a Situation orientation was discarded for having a regression coefficient that was more than 3 standard deviations above the grand mean.

Using the remaining 75 participants, a one-way ANOVA with Orientation as a between-subjects factor revealed no differences in the regression coefficients between levels of Orientation, $F(2, 72) = 0.30$, $p = .75$, $\eta^2_{partial} = .01$. Consequently, data were collapsed across levels of Orientation and collectively compared to zero using a one-sample *t*-test. The mean regression coefficient (M = -0.13, SE = 0.07) was less than zero but not significantly so, $t(74) = -1.72$, $p = .089$, $d = -0.20$. Thus, in Experiments 1 and 2, we obtained evidence for a negative relationship between confidence in one's ability to predict the future and the likelihood of reporting an event boundary in the near future. However, this finding was not statistically reliable in Experiment 2.

3.2.4 Discriminability

Log *d* values were analyzed with a 3 x 3 mixed ANOVA with Probe Type (Within/Within, Within/Across, Across/Across) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Probe Type, $F(2, 162) = 8.23$, $p < .001$, $\eta^2_{partial} = .09$, but no significant effect of Orientation, $F(2, 81) = 1.06$, $p = .35$, $\eta^2_{partial} = .03$, nor interaction between Probe Type and Orientation, $F(4, 162) = 0.72$, $p = .58$, $\eta^2_{partial} = .02$.

Trend analysis of Probe Type indicated a significant linear, $F(1, 81) = 16.84$, $p < .001$, $\eta^2_{partial} = .17$, but not a significant quadratic trend, $F(1, 81) = 1.26$, $p = .27$, $\eta^2_{partial} = .02$. Thus, the ability to discriminate between targets and foils was greatest when within a situation and fine event (Within/Within; M = 0.22, SE = 0.03), worse when nearing the end of a fine event during an ongoing situation (Within/Across; M = 0.18, SE = 0.03), and even worse when nearing the end of both a fine event and situation (Across/Across; M = 0.05, SE = 0.03; Figure 10).

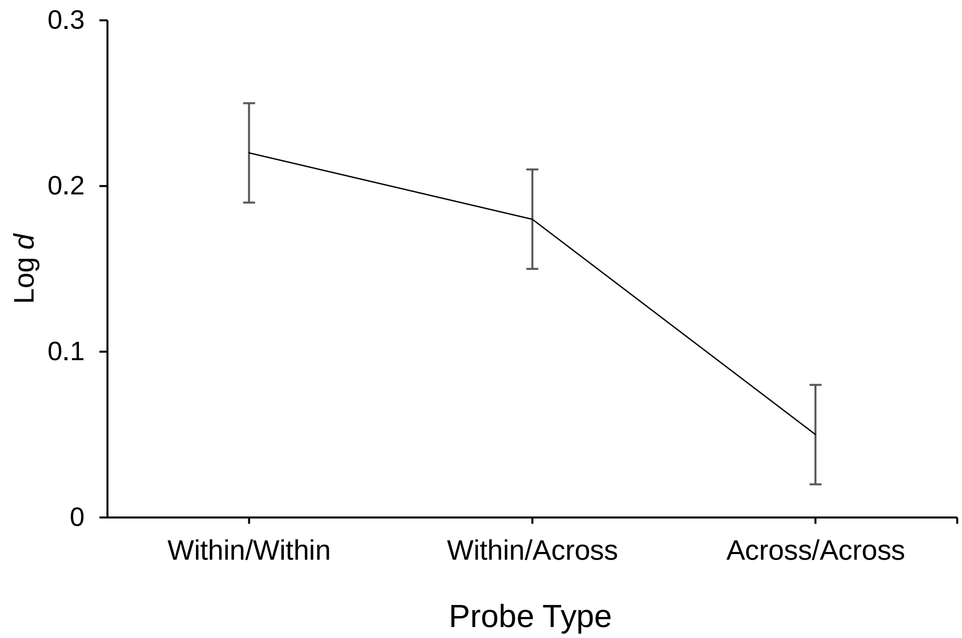

Figure 10 – Log *d*, a measure of discriminability, as a function of Probe Type. Error bars are the standard error of the mean.

3.2.5 Reaction Time

For each participant, we computed the mean reaction time of correct responses for each level of Probe Type. We excluded the data for one participant who did not make any correct responses to Within/Across probes. Mean reaction times were subjected to a 3 x 3 mixed ANOVA with Probe Type (Within/Within, Within/Across, Across/Across) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Probe Type, $F(1.84, 146.9) = 13.55$, $p < .001$, $\eta^2_{partial} = .15$, but no significant effect of Orientation, $F(2, 80) = 1.84$, $p = .17$, $\eta^2_{partial} = .04$, nor interaction between Probe Type and Orientation, $F(3.67, 146.9) = 0.08$, $p = .99$, $\eta^2_{partial} = .002$.

Trend analysis of Probe Type indicated a significant quadratic trend, $F(1, 80) = 21.22$, $p < .001$, $\eta^2_{partial} = .21$, but not a linear trend, $F(1, 80) = 0.19$, $p = .66$, $\eta^2_{partial} = .002$. The quadratic trend suggests that reaction time to a probe is slower when nearing the end of a fine event in a situation (Within/Across; $M = 3.42$ seconds, $SE = 0.22$) than when within a fine event and a situation (Within/Within; $M = 2.42$, $SE = 0.15$) or nearing the end of a fine event and a situation (Across/Across; $M = 2.50$, $SE = 0.14$; Figure 11).

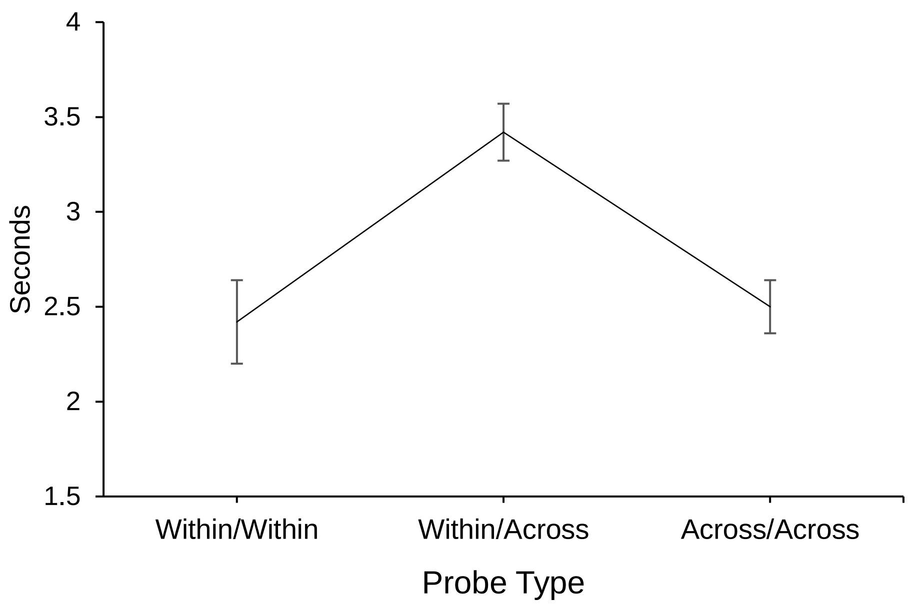

Figure 11 – Mean reaction time for correct responses to probes as a function of Probe Type. Error bars are the standard error of the mean.

3.2.6 Response Bias

Log *b* values were subjected to a 3 x 3 mixed ANOVA with Probe Type (Within/Within, Within/Across, Across/Across) as a within-subjects factor and Orientation

(Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Probe Type, $F(1.82, 147) = 23.26$, $p < .001$, $\eta^2_{partial} = .22$, but no significant effect of Orientation, $F(2, 81) = 0.55$, $p = .58$, $\eta^2_{partial} = .01$, nor interaction between Probe Type and Orientation, $F(3.63, 147) = 1.74$, $p = .15$, $\eta^2_{partial} = .04$.

Trend analysis of Probe Type indicated a significant linear, $F(1, 81) = 35.25$, $p < .001$, $\eta^2_{partial} = .30$, but not a significant quadratic trend, $F(1, 81) = 0.01$, $p = .94$, $\eta^2_{partial} < .001$. The tendency to respond "yes" to probes was strongest when within a situation and fine-grained event (Within/Within; $M = -0.36$, $SE = 0.03$), weaker when nearing the end of a fine-grained event in a situation (Within/Across; $M = -0.20$, $SE = 0.03$), and even weaker when nearing the end of both a fine-grained event and situation (Across/Across; $M = -0.05$, $SE = 0.04$; Figure 12). In Appendix B, the interested reader can find comparable analyses of untransformed hit rates and false alarm rates.

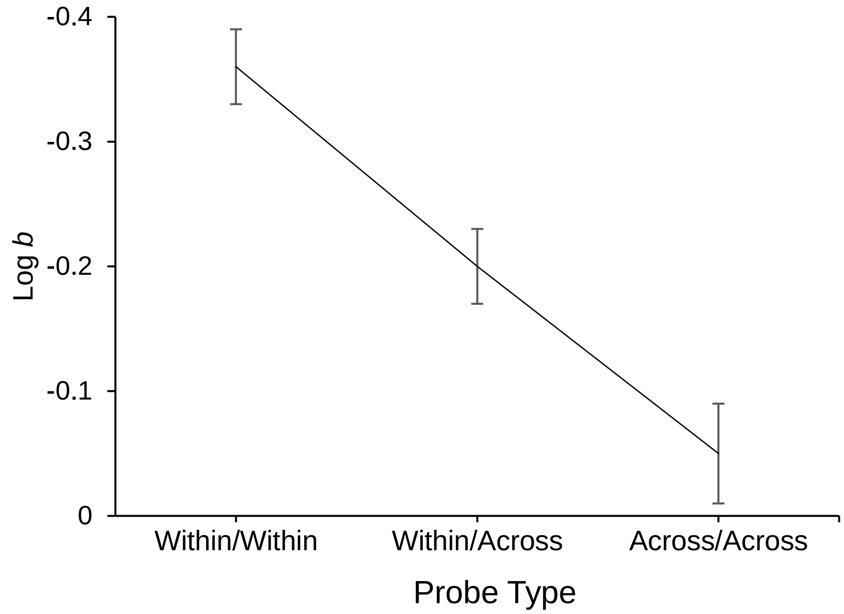

Figure 12 – Log *b*, a measure of response bias, as a function of Probe Type. Error bars are the standard error of the mean.

3.3 Discussion

Like Experiment 1, Experiment 2 provided some evidence for the general mechanisms of EST but failed to support the hypothesis that the hierarchical structure of event segmentation is attributable to the differential predictive accuracies of event models on different timescales. Regarding the former, we observed that one's confidence in their ability to predict what happens next was negatively related to the likelihood of perceiving an event boundary in the near future (although not significantly so, p = .089). Regarding the latter, both the confidence in and the accuracy of one's prediction varied as a function of Probe Type only.

Unlike Experiment 1, however, confidence and accuracy were not affected by the levels of Probe Type in the same way. Confidence was highest during a fine event in an ongoing situation (Within/Within) and was equally low near the end of a fine event in an ongoing situation (Within/Across) and near the end of a fine event and the current situation (Across/Across). This replicates the effect of Probe Type on confidence ratings observed in Experiment 1 and is consistent with our expectations for fine segmentation. Predictive accuracy, on the other hand, was highest within a fine event in an ongoing situation (Within/Within), worse when nearing the end of a fine event in an ongoing situation (Within/Across), and even worse when nearing the end of both a fine event and situation (Across/Across). Given that accuracy and confidence diverged only at Within/Across probes, which participants took longer to answer correctly than either Within/Within or Across/Across probes, it is possible that participants were able to deduce the correct answer. This reveals a potential weakness of using Yes/No decision tasks to assess predictive accuracy when comprehending ongoing activity (e.g., Huff et al., 2014; Zacks et al., 2011). At the same time, it might also suggest that when given a possible future, one's current fine-event model can be used to infer the next fine event in an ongoing situation, but *not* when the next fine event is part of a new situation.

Lastly, we found that response bias was most liberal during a fine event in an ongoing situation (Within/Within), less so when nearing the end of a fine event in an ongoing situation (Within/Across), and even less so when nearing the end of a fine event and the current situation (Across/Across). Although we did not hypothesize how response bias might be affected by the levels of Probe Type, this result is similar to one obtained by Zacks et al. (2011), who found that response bias was more liberal within an event than near the end of an event.

Like Experiment 1, the failure to observe an effect of Orientation on confidence or accuracy cannot be attributed to a failure to manipulate Orientation. Situation and fine segmentation were differentially related to the normative situation and fine segmentation of a different group of participants who segmented *The Red Balloon*. Moreover, situations (and neutral-grain events) were segmented significantly less frequently than were fine events throughout the entire film.

Otherwise, neutral-grain segmentation in Experiment 2 was indistinguishable from situation segmentation, which contradicts our earlier finding that participants may have preferentially segmented the film into small events. Given that Experiments 1 and 2 only differed with respect to the prediction task, this shift in segmentation grain may have been a response to task demands, to which neutral-grain segmentation is sensitive. For example, Cohen and Ebbesen (1979) had participants freely segment a video of an actor performing a task with either the goal of forming an impression of the actor's personality or learning the actor's task. Participants who segmented to form an impression of the actor produced larger segments than did participants who segmented to learn the task. Lassiter, Geers, Apple, and Beers (2000), however, found that having either goal (i.e., forming an impression or learning the task) resulted in coarser-grained segmentation than having no specific goal during segmentation and found no difference in segmentation rate between participants with either goal.

Like Experiment 1, the results of Experiment 2 also suggest that participants had event models of activity only at the fine-grained level, despite the fact that segmentation clearly varied as a function of Orientation. In support, the effect of Probe Type on confidence ratings and predictive accuracy were consistent with that expected for fine segmentation.

CHAPTER 4. GENERAL DISCUSSION

To behave adaptively in our everyday lives, it is crucial that we recognize when the current situation ends and a new situation begins. Central to this imperative is the issue of how we form and modify our internal representations of situations, which are one of the fundamental mediating mechanisms between external situations and psychological outcomes (Parrigon et al., 2017, p. 670). A review of several related but largely unintegrated literatures (e.g., social psychology, engineering psychology, and narrative comprehension) suggests that our representations of situations are bounded, hierarchical, multifaceted, and future-oriented (at least for experienced situations). We argued that each of these properties is embodied by Event Segmentation Theory (EST; Zacks et al., 2007) and moreover, that EST can provide a novel framework for understanding how these properties coalesce to give rise to our representations of situations.

In the present studies, we were concerned with how EST might account for one of these properties in particular - the hierarchical structure of our representations of situations. EST also claims that our representation of ongoing activity is hierarchical, such that people simultaneously perceive a series of fine events unfolding within coarse events. Specifically, people maintain internal models of events unfolding on different timescales simultaneously in working memory. Each “event model” strives to predict the state of the world in the near future and will continue to represent ongoing activity as long as its predictions are consistent with sensory input. However, when large discrepancies between predictions and reality emerge, the current event model is updated, giving rise to the experience of a perceptual boundary between events.

According to EST, the hierarchy of event models in working memory arises because event models maintained on separate timescales can differ in the amount of prediction error they experience; specifically, when a fine event boundary occurs within the boundaries of a coarse event, the coarse event model should experience less prediction error in that moment than the fine event model. We directly tested this hypothesis by orienting participants to event models on either the situation (i.e., coarse) or fine level of activity in a film and assessed their confidence and predictive accuracy at moments when both variables should depend upon the event model being interrogated. These moments were either during a fine event in an ongoing situation (Within/Within), when nearing the end of a fine event in an ongoing situation (Within/Across), or when nearing the end of both a fine event and situation (Across/Across).

Because the present studies were the first to simultaneously assess confidence, predictive accuracy, and overt segmentation, we also tested several critical but unexamined hypotheses regarding these three variables. Specifically, EST claims that event models are updated in response to spikes in prediction error, with the magnitude of prediction error being related to one's subjective uncertainty about the future (Zacks et al., 2011). However, no study has tested whether confidence and predictive accuracy are predictive of subsequent segmentation.

In Experiment 1, we found that the more confident one was in their ability to predict what would happen next in the film, the more consistent his or her prediction was with what actually happened in the near future. Furthermore, the more consistent one's prediction was with what happened next in the film, the less likely he or she was to perceive an event boundary in the near future. In both Experiments 1 and 2, we observed that confidence in one's ability to predict what would happen next in a narrative film was negatively related to the likelihood of perceiving an event boundary in the near future (although this effect was not reliable in Experiment 2). Thus, we obtained novel and, in one case, somewhat replicable evidence for the general mechanisms of EST.

Although we found support for the general mechanisms of EST, it appears that these mechanisms were operating only in event models of activity at the fine level. In Experiments 1 and 2, confidence in one's ability to predict what would happen next followed a pattern expected for fine segmentation. Specifically, participants were most confident in their ability to predict the future during a fine event in an ongoing situation (Within/Within) but were equally lower in confidence either when nearing the end of a fine event in an ongoing situation (Within/Across) or when nearing the end of a fine event and a situation (Across/Across). This pattern suggests that confidence was varying as a function of whether a fine event boundary was imminent or not.

Regarding predictive accuracy, we used complementary operationalizations (i.e., the "three-prong method"; Magliano & Graesser, 1991) to ensure that our conclusions would not be endemic to one approach. Indeed, our two measures of predictive accuracy largely converged on the same pattern that, like confidence, was consistent with fine segmentation; Participants were most accurate at predicting what happened next in the film during a fine event in a situation (Within/Within) and were less accurate when nearing the end of a fine event and situation (Across/Across). As Kurby & Zacks (2012) observe, verbal protocols and behavioral measures (e.g., reading time) often correspond to one another, yet when they do diverge, it is because the former is less sensitive to subtle changes in event structure than the latter. Indeed, when nearing the end of a fine event during an ongoing situation (Within/Across), our two measures of predictive accuracy diverged. In Experiment 1, predictive accuracy (i.e., consistency) was just as poor at these moments as it was when nearing the end of a fine event and situation (Across/Across), replicating the pattern observed for confidence in both Experiments 1 and 2. In Experiment 2, predictive accuracy (i.e., discriminability) at these moments was worse than during a fine event in a situation (Within/Within) but better than when nearing the end of a fine event and situation (Across/Across). Taken together, our two measures of predictive accuracy suggest that people are less likely to know what happens next when nearing the end of a fine event than when within a fine event (Experiment 1). However, when given a possible future (i.e., a target or foil; Experiment 2), people are able to identify the correct answer with a moderate degree of accuracy, but *only* if the next fine event is part of same ongoing situation. This might suggest that one's current fine-event model can be used to infer the next fine event in an ongoing situation, which was only intimated when we assessed predictive accuracy using a Yes/No decision task.

The present studies replicate previous observations that both confidence (Zacks et al., 2011) and predictive accuracy (Eisenberg et al., 2018; Reynolds et al., 2007; Zacks et al., 2011) are higher within an event than at the end of an event. Critically, the present studies suggest that this may only be true for our representations of fine events, but not coarse events (e.g., situations). Despite evidence that participants only had fine event models in the present studies, however, their overt segmentation clearly depended on their orientation to activity in the film. Across Experiments 1 and 2, we consistently found that situation segmentation was related only to the normative situation segmentation of a separate group of participants whereas fine segmentation was related to both normative fine and situation segmentation. Moreover, participants who were oriented to situations segmented the film more coarsely than did participants oriented to fine activity (although this finding was not reliable in Experiment 1).

Interestingly, the grain at which people naturally segmented the film (i.e., neutral-grain segmentation) differed between the two studies; in Experiment 1, neutral-grain segmentation was similar to fine segmentation, but in Experiment 2, neutral-grain segmentation was indistinguishable from situation segmentation. That the preferred granularity of segmentation differed systematically between the two studies suggests that the natural grain may have been dictated by task demands (e.g., Cohen & Ebbesen, 1979; Lassiter et al., 2000), although it is not obvious why one prediction task would engender fine but not situation-level segmentation (and vice versa). Importantly, whether participants naturally oriented to the fine or situation level of activity was unrelated to their confidence and predictive accuracy; in both cases, their performance was consistent with having fine-event models.

We observed a divergence between how observers segment ongoing activity and the event models of activity that observers maintain. Specifically, the present studies found no evidence that people maintain coarse-event models of situations; regardless of how one was oriented to activity in the film, his or her confidence and predictive accuracy depended entirely on the structure of fine events. This observation conflicts with a premise of EST, which is that event boundaries identified during segmentation correspond to an event model being updated, with fine boundaries marking the advent of a new fine-event model and coarse boundaries marking the advent of a new coarse-event model. Like EST, we believe that fine segmentation may indeed reflect the updating of fine-event models, but unlike EST, we propose that coarse segmentation may instead reflect how observers group fine events online rather than the updating of coarse event models per se (Zacks et al, 2007, p. 286). Accordingly, we argue that when people overtly segment activity into situations, they are reporting each time they perceive that one meaningful sequence of fine events (e.g., an episode) has ended and a new sequence has begun.

We argue that our interpretation is a more parsimonious view of the hierarchical structure of event segmentation than that of EST. EST proposes that people maintain a hierarchy of event models simultaneously in working memory (Radvansky & Zacks, 2011; Radvansky & Zacks, 2014; Zacks et al., 2007). On any given timescale, there is only one active event model, but at any given moment, there are multiple active event models on different timescales (Radvansky & Zacks, 2014). Yet, working memory is limited in how much information it can maintain actively and how long it can maintain that information (Baddeley, 2003). Thus, maintaining multiple event models might incur a considerable cost given how difficult it is to predict the future (e.g., Wickens, 2015), let alone make *n* (perhaps as many as six; Radvansky & Zacks, 2014) simultaneous predictions for timescales ranging from seconds to tens of minutes. Again, we advocate for a simpler view – we do maintain event models of finer units of activity and these models attempt to predict the very near future. Rather than perceiving an endless succession of fine events, however, we actively organize a sequence of fine events in terms of a single coherent situation.

Our interpretation is also compatible with broad evidence that people tend to perceive coarse events as being composed of fine events. We have observed that the boundaries people tend to identify between situations per se are a subset of the boundaries that people tend to identify between fine events (Mumma & Durso, in revision), consistent with evidence that coarse-event boundaries tend to be a subset of fine-event boundaries, regardless of whether segmentation grain is varied between observers (e.g., Hanson & Hirst, 1989; Newtson, 1973) or within observers (e.g., Hard et al., 2006; Zacks et al., 2001). Indeed, the temporal alignment of coarse and fine boundaries is a highly replicable characteristic of event segmentation (e.g., Hard et al., 2006; Kurby & Zacks, 2011; Kurby & Zacks, 2012; Sargent et al., 2013; Zacks, 2004; Zacks et al., 2001; Zacks et al., 2006; Zacks et al., 2009; Zacks et al., 2010) and, in some cases, reflects the attainment of higher level goals and lower level goals, respectively (Kurby & Zacks, 2019; Zacks et al., 2001). Moreover, the boundaries of coarse events also tend to "enclose" fine events (Hard et al., 2006; Hard et al., 2011; Zacks et al., 2009). That is, coarse boundaries tend to fall just slightly after the nearest fine unit boundary in time, which suggests that fine events are subsumed by coarse events (Kurby & Zacks, 2008).

If coarse event boundaries mark the end of one group of related fine events and the start of a new group, it is reasonable to expect that processing coarse boundaries should be more difficult than fine boundaries. Converging behavioral and neurophysiological evidence supports this hypothesis. As discussed earlier, brain responses are generally larger at coarse than fine event boundaries (Kurby & Zacks, 2018; Speer et al., 2007; Zacks et al., 2001; Zacks et al., 2006), which suggests that coarse boundaries engender greater processing demands than do fine boundaries (Zacks et al., 2001). Using electrophysiological techniques, Delogu, Drenhaus, and Corcker (2018) found that coarse boundaries, which marked the start of a new unrelated activity in a story (e.g., jogging versus dishwashing), elicited a stronger N400 (reflecting the retrieval of lexical semantic information) and P600 (reflecting situation model updating) response than did fine boundaries, which marked the start of a new event that continued the current activity in a story (e.g., drying dishes after washing dishes).

That coarse boundaries are more difficult to process than fine boundaries is also evidenced by behavioral data, such as dwell times. When advancing through a flip book depicting an everyday activity, people look longest at coarse boundaries, less at intermediate boundaries, and even less at fine boundaries (Hard et al., 2011; Kosie & Baldwin, 2019). Moreover, this effect can only be partially explained by the magnitude of low-level physical changes occurring at boundaries, which suggests that conceptual changes may also contribute to dwell time above and beyond perceptual changes occurring at boundaries. In the present studies, predictive accuracy near the end of a fine event and situation (Across/Across) similarly suggests that processing the end of a situation is more difficult than processing the end of a fine event in an ongoing situation. Although Experiment 1 revealed that people are less likely to know what will happen in the next fine event than later in the current fine event, they are better at identifying what happens in the next fine event as long as it is part of the same situation; if the next fine event is part of a new situation, however, predictive accuracy drops to near chance levels.

Thinking of coarse grain (i.e., situation) segmentation as the online chunking of fine event models suggests an important role for event knowledge (e.g., knowing what events precede or follow one another or are part of achieving a larger goal). When participants can choose their preferred grain of segmentation, having knowledge about an activity or actor tends to engender coarser-grained segmentation (e.g. Bläsing, 2015; Graziano et al., 1988; Levine et al., 2017; Markus et al., 1985; Massad et al., 1979; Newberry & Bailey, 2019; Newtson, 1973; Wilder, 1978a, Wilder, 1978b). Moreover, Hard, Meyer, and Baldwin (2018) demonstrated that the perception of coarse but not fine boundaries (Hard et al., 2011; Kosie & Baldwin, 2019) emerges after learning the higher-level structure of activity. Hard et al. (2018) created coarse events or "actions" by randomly selecting and ordering three "small motion events" (e.g., stacking, poking, or drinking) from a larger corpora of such events. One group of participants was exposed to the statistical regularities of small motion events (i.e., how they form larger actions) while another other group remained naïve. Later, participants advanced through a flipbook of actions at their own pace while their dwell times were recorded. All participants looked longer at slides depicting boundaries between small motion events than slides depicting within-event content. However, only participants who had acquired knowledge of the statistical regularities of small motion events slowed down at boundaries separating actions (i.e., recurring patterns of small motion events).

Lastly, if grouping a series of fine events into a coarse event depends upon event knowledge, then reducing the availability of this knowledge should selectively impair coarse but not fine grain segmentation. Converging neuropsychological evidence supports this hypothesis. For example, Zalla, Labruyère, and Georgieff (2013) observed that individuals with mental retardation or learning disabilities were significantly worse than typically developing controls at identifying prototypical coarse but not fine boundaries, owing to the reduced effect of conceptual knowledge during online segmentation. Moreover, patients with psychiatric (e.g., schizophrenia; Zalla, Verlut, Franck, Puzenat, & Sirigu, 2004) or neurologic disorders (e.g., frontal lobe damage; Zacks, Kurby, Landazabal, Kruger, & Grafman, 2016; Zalla, Pradat-Diehl, & Sirigu, 2003) associated with frontal cortex dysfunction are also selectively impaired at identifying coarse but not fine boundaries, relative to normal controls. Notably, the frontal cortex contains subregions that represent knowledge of social events and routine event sequences (Zacks et al., 2016) and are also important for planning and executing complex behavior (Zalla et al., 2003).

4.1 Limitations

Because the present studies used a narrative feature length film, there are potential limits on the generalizability of our findings. For example, filmmakers use certain conventions to tell stories (e.g., cuts, framing, or music; Zacks & Magliano, 2011), which are typically not part of the situations we encounter in everyday life. Thus, how participants understood the situations in *The Red Balloon* may reflect the influence of these conventions (e.g., because of how they can powerfully direct attention; Magliano et al., 2014). However, there is evidence that we form representations of situations while observing less-scripted activities (e.g., a first-person shooter game; Magliano et al., 2014) or participating in interactive events (e.g., in actual or virtual reality environments; Radvansky & Copeland, 2006; Radvansky, Krawietz, & Tamplin, 2011) that are similar to those formed in more structured contexts (e.g., reading or watching a narrative film; Magliano et al., 2014).

Furthermore, participants in the present study attempted to understand another person's situations, rather than their own personal situation. Ultimately, our goal is to understand the latter, although we suspect there is much overlap between the processes underlying each. For example, people are accurate at perceiving the situations of others; when rating the characteristics of a situation based on limited information about that situation (e.g., who or what was present and where the situation occurred), there is considerable agreement between raters *in situ* (i.e., people rating their own description of a situation they recently experienced) and raters *ex situ* (i.e., people rating the descriptions of situations provided by raters *in situ*; Sherman et al., 2010; Sherman, Nave, & Funder, 2013; Rauthmann et al., 2014; Rauthmann & Sherman, 2017). Moreover, people tend to identify the same coarse and fine event boundaries when segmenting an activity filmed from either the first or third-person perspective (Swallow et al., 2018). Thus, observers arrive at similar representations of situations (e.g., their characteristics and how they unfold) regardless if they are experiencing or observing a situation.

Furthermore, people are often active participants in situations, rather than passive observers like participants in the present studies. Therefore, one might doubt whether knowing how observers understand ongoing activity translates to how people understand the activities of which they are a part. Although there is a lack of research on event model construction during live events (Richmond & Zacks, 2017), there appears to be overlap between the systems for understanding activity (i.e., event cognition) and the systems for producing activity (Bailey et al., 2013) and perhaps even a common representational format (Hommel et al., 2001; Kurby & Zacks, 2008).

Lastly, we assessed predictive accuracy in both studies using probes, which temporarily halted the movie to assess a participant's confidence and predictive accuracy. There are theoretical reasons to believe that probes may have been disruptive (i.e., as interruptions; Foroughi, Werner, Barragán, & Boehm-Davis, 2015). Indeed, we observed data suggesting that the effect of orientation on overt segmentation was minimal shortly after probes, but our global manipulation checks strongly suggested that this effect was temporary. Nonetheless, less invasive assessments of predictive accuracy (e.g., via eye-movements; Eisenberg & Zacks, 2016) may be preferable in future research so as not to disrupt segmentation.

4.2 Conclusions and Future Directions

In summary, the present studies suggest that situation (i.e., coarse) segmentation may reflect how observers chunk a series of fine events together into a single situation whereas fine segmentation may reflect the updating of fine-event models. Not only is this interpretation consistent with broad evidence that our perceptions of coarse events are built from fine events but it is also consistent with the view of situations as episodes of activity, which is found across different literatures that study the representations people form of situations (e.g., Deutsch et al., 1994; Flach et al., 2004; Forgas, 1976; Radvansky & Wyer, 1999). Lastly, the findings from the present studies suggest future directions for both application and theory.

4.2.1 Applied Directions

Regarding the practical implications of the present studies, we suggest that event segmentation be assessed in operators of dynamic environments, given the importance of tracking and connecting events over time to their situation awareness (Durso et al., 2007; Sarter & Woods, 1999; Woods, 1988). For example, Christoffersen et al. (2007) used a modified version of the event marking paradigm to elicit the meaningful events that anesthesiologists perceived during a simulated telemetry monitoring task. In dynamic environments, however, it is particularly important to understand how operators segment ongoing activity in real time, rather than in a video playback. Several studies suggest that the hierarchical segmentation of ongoing activity may be evinced by eye-movements (Eisenberg & Zacks, 2016; Hard et al., 2011; Kosie & Baldwin, 2019), which could provide a less invasive measure of event segmentation than the traditional event marking paradigm (Newtson, 1973).

Moreover, not only would assessing event segmentation be important for understanding an individual's situation awareness, but also the collective situation awareness of a team of operators. Several studies have shown that certain psychological outcomes (e.g., memory and the efficiency of action execution; Bailey et al., 2013; Kurby

& Zacks, 2011; Zacks et al., 2006) are related to how normatively one segments an activity, suggesting that normative segmentation is related to efficient cognitive functioning. In the context of a team managing a situation (e.g., medical responders resuscitating a patient in cardiac arrest), it is plausible that the extent to which the online segmentation of different teammates is coordinated would be positively related to team outcomes (e.g., performance).

4.2.2 Theoretical Directions

Regarding future theoretical directions, the present studies raise at least three important questions. The first is that if situations are indeed perceived as chunks of fine events, then what mechanism(s) may underly this chunking? Our previous discussion of the relationship between coarse and fine events strongly suggests that event knowledge contributes to how observers group fine events into coarse events. More specifically, several studies suggest that knowledge about the goal and sub-goal relationships between events can give rise to the hierarchical structure of event segmentation (e.g., Kurby & Zacks, 2018; Zacks et al., 2001), although such knowledge is not always necessary (Hard et al., 2006). That the knowledge of goals and sub-goals may structure our representation of activity resonates with a recent conceptualization of situations in the social psychology literature. In a comprehensive review of the definitions of situations, taxonomies of situations, and interrelationships of situations, people, and behavior, Yang, Read, and Miller (2009, p. 1018) concluded that "…the essence of a situation is its affordance of human goals, and that situations are largely characterized by two specific principles of goal processes (what happened, is happening, or might happen to people's goals) and goal contents (the specific goals afforded in the situation).” Inferring the goals that are afforded by a situation (and the behaviors of people in the situation that are related to those goals; Magliano, Skowronski, Britt, Güss, & Forsythe, 2008; Rauthmann, 2016) could be a powerful mechanism for organizing a coherent sequence of fine events. Thus, knowledge of goals may facilitate the organization of fine events into coarse events in general, but also into situations specifically.

The second question concerns whether our representations of fine and coarse events interact with each other; specifically, are fine events simply atomic or primitive things (see Zacks & Tversky, 2001, p. 4) that we group into larger events or do fine events themselves depend on how one is chunking activity? That coarse grain, but not fine grain segmentation, can be selectively impaired would seem to support the idea that fine events can be perceived independently of coarse events (e.g., Zacks et al., 2016; Zalla et al., 2003; Zalla et al., 2004; Zalla et al., 2013). Moreover, the boundaries between fine events are more strongly related to lower-level perceptual changes (e.g., in the direction or speed of an object or actor) than are the boundaries between coarse events (Hard et al., 2006; Newtson, et al., 1977; Zacks, 2004; Zacks et al., 2009), suggesting that fine events may also have a less conceptual basis than coarse events. However, in domains such as aviation or driving, an inexperienced operator placed in the same situation as an experienced operator may not perceive an event whereas the experienced operator does (e.g., a hazardous event; Durso, Dattel, & Pop, 2018). This would suggest that fine events may not be perceptual primitives but rather, do sometimes require knowledge to be perceived.

A final question is whether people ever maintain multiple event models simultaneously. It would be premature to conclude that we only have fine-event models of activity from the present two studies alone. One possibility is that our findings are limited to the kind activity portrayed in *The Red Balloon* – extended, unfamiliar, naturalistic activity. It remains possible that in highly structured (e.g., ritualistic) or routine situations, individuals do maintain fine- and coarse-event models simultaneously because prediction on more than one timescale is tenable. Therefore, future research should strive to uncover the conditions under which people do maintain multiple event models of activity simultaneously, if any.

APPENDIX A: PREDICTIONS AT PROBE LOCATIONS

Table 5 – Predictions and their frequency (*f*) at Probe 1 (Within/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy pets the cat, then walks away.	7	The boy continues walking.	6	The boy and the cat walk off together.	12
The boy continues walking.	6	The boy and the cat walk off together.	5	The boy continues walking.	6
The boy continues interacting with the cat.	5	The boy continues interacting with the cat.	3	The boy continues interacting with the cat.	3
A new person appears.	4	The boy pets the cat, then walks away.	3	The cat runs away.	3
The cat follows the boy.	4	The cat follows the boy.	3	A new person appears.	2
The cat runs away.	3	The cat runs away.	3	The child walks out of frame.	1
The boy will pet the cat and the camera will shift focus onto a new character.	1	The boy walks home.	2	The boy is going to a formal occasion.	1
--	--	Someone else appears on screen.	1	Transition from this scene to another.	1
--	--	The child wanders the street and his/her parents go looking for him/her.	1	He will go home.	1
--	--	The same activity continues until a new one comes up.	1	--	--
--	--	Someone will come to pick up the child and the dog.	1	--	--
--	--	The main character gets introduced?	1	--	--

Table 6 – Predictions and their frequency (*f*) at Probe 2 (Within/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy continues walking with the balloon.	24	The boy continues walking with the balloon.	21	The boy continues walking with the balloon.	16
The boy loses the balloon.	2	The boy takes the balloon home.	4	The boy continues walking.	4
The boy walks home with the balloon.	2	The boy loses the balloon.	3	The boy goes back to the cat.	3
The boy takes the red balloon with him to school.	1	He will eventually meet others who will try to take the balloon from him as he will try to recover the balloon.	1	The boy gives his balloon to someone else.	2
The boy is going to play with the balloon.	1	The child lands on the ground.	1	Someone else approaches the child.	1
--	--	--	--	He will hold the red balloon and go to his house or back where he came from.	1
--	--	--	--	The boy will get down and drop his balloon.	1
--	--	--	--	The boy will land with the balloon which might float away.	1
--	--	--	--	The boy will find his parents.	1

Table 7 – Predictions and their frequency (*f*) at Probe 3 (Across/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy catches up to the bus.	8	The boy chases after the bus.	10	The boy catches up to the bus.	7
The boy chases after the bus.	8	The boy runs home.	3	The boy runs to his destination.	6
The boy keeps running.	3	The boy loses the balloon.	2	The boy chases after the bus.	5
The boy does not catch the bus.	2	The boy runs to his destination.	2	The boy does not catch the bus.	3
The boy reaches his destination.	2	The boy tries to catch another bus.	2	The boy runs home.	3
He runs home with his balloon.	1	The child reaches one of his destinations.	1	The boy loses the balloon.	2
He will get tired and give up.	1	The boy continues running down the sidewalk.	1	The boy runs into something or gets hit.	1
Boy rounds a corner.	1	The boy would be able to catch up the transport and get it on.	1	Child runs so fast the balloon begins to make him float.	1
The kid falls/gets tired/gets injured.	1	He runs until he gets tired.	1	The boys running causes him to upset people and causes a scene in which the balloon pops.	1
The boy runs to a door and enters the building.	1	Boy stops in a store/shop.	1	The boy goes into a building.	1
The boy will find some form of alternative transportation.	1	The kid doesn't catch the bus and he starts to walk to school.	1	--	--
Somebody helps the boy.	1	The boy fails to catch the bus.	1	--	--
--	--	The child will get hurt in some way.	1	--	--
--	--	He turns on a corner or runs into someone.	1	--	--
--	--	The boy encounters something (the bus he missed, a car, person, etc.).	1	--	----
--	--	The boy stops running.	1	--	--

Table 8 – Predictions and their frequency (*f*) at Probe 4 (Within/Within) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The man walks the boy home.	5	The man and the boy continue walking together.	7	The man and boy part ways.	5
The boy walks himself home.	4	The man walks the boy home.	5	The boy walks himself home.	4
The man and the boy continue walking together.	3	The man and boy part ways.	2	The man and the boy walk into a building.	3
The balloon will pop.	2	The man and the boy stop to get food.	2	The man and the boy walk into a store.	3
The man and boy part ways.	2	The man and the boy walk into a store.	2	The balloon will pop.	2
The man and the boy stop to get food.	2	The man and boy will become friends.	1	The boy goes to the man's house.	2
The man and the boy walk into a building.	2	Gentleman with the booklet will follow him.	1	The man walks the boy home.	2
The man and the boy walk into a store.	2	The boy and man are going to find somewhere inside to go.	1	He continues to look for someone or get help since he is lost.	1
He goes to the old man's house.	1	The boy again left the balloon to the old person.	1	The boy keeps finding others to keep his red balloon.	1
The man from the window will follow the boy and his new friend.	1	They will walk together, but the balloon may be taken away by wind.	1	The man invites the kid to go with him.	1
The boy will befriend a stranger with his balloon.	1	We see the man observing him again.	1	The child is approached by another stranger.	1
The man has to leave and leaves little boy w/o an umbrella.	1	The boy will go to the old man with umbrella's house.	1	The boy is sending the balloon to an important person that he loves.	1
The boy finds his family.	1	The boy and man will sit and talk.	1	The boy will hand the man the balloon and leave.	1

Table 8 (continued).

He'll say goodbye & run with his balloon to a building (home? a store?).	1	The man points somewhere indicating where he or the kid should go.	1	The boy and man will continue their conversation and hatch a plan so the boy can keep his balloon for the ride.	1
The boy gets on the bike.	1	The kid will thank the man for letting him under the umbrella and reach his destination.	1	The man who was taking notes will confront the boy.	1
--	--	The boy will find a way to keep the balloon with him all day long while doing all the other normal things he does every day.	1	The boy will lose the balloon.	1
--	--	The child starts climbing the stairs again.	1	--	--

Table 9 – Predictions and their frequency (*f*) at Probe 5 (Across/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy releases the balloon outside.	8	The boy releases the balloon outside.	11	The boy releases the balloon outside.	7
The boy goes outside with his balloon.	5	The boy goes outside with his balloon.	7	The boy brings the balloon inside	4
The boy looks down onto the street.	4	The boy looks down onto the street.	4	The boy goes outside with his balloon.	4
The boy loses the balloon.	2	Kid ties balloon outside window.	1	The boy gives the balloon away to someone.	2
The boy runs.	2	The boy will try to talk to the street vendor.	1	The boy goes outside.	2
The boy takes the balloon to school.	2	The lady reads the letter.	1	The boy interacts with the mailman.	2
The boy walks out onto the balcony.	2	The child loses his balloon.	1	Mailman delivers more mail to people.	1
The boy yells out to the street seller.	1	Plays with balloon.	1	The little boy's balloon goes down to get whatever the man's selling.	1
The boy and the balloon are going to explore his house.	1	The mailman continues to handout mail.	1	The balloon explores the city.	1
The boy will hide the balloo n somewhere.	1	The child will go & play with his balloon some more but something will happen to him/it.	1	The boy ties the balloon to the iron bars outside to prevent it from being seen by people in the room.	1
The boy will tie the balloon up & then go on another outing.	1	Someone (e.g. the mailman) will come knock on the boy's door.	1	The child will try to get rid of the balloon in some other way.	1
The boy hears the man and meets him.	1			The kid brings the balloon back to his house and his mother will notice it.	1

Table 9 (continued).

--	--	--	--	The woman catches the kid with the balloon and scolds him.	1
--	--	--	--	The kid will let the balloon see the sun rise with him.	1
--	--	--	--	The boy will interact with someone outside his window.	1

Table 10 – Predictions and their frequency (*f*) at Probe 6 (Within/Within) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy is not allowed on the bus.	22	The boy is not allowed on the bus.	19	The boy is not allowed on the bus.	20
The balloon follows the boy as he rides the bus.	4	The balloon follows the boy as he rides the bus.	6	The balloon follows the boy as he rides the bus.	4
The boy gets on the bus with the balloon.	1	The boy boards the bus.	3	The boy boards the bus.	2
He is going to try to get on the bus.	1	The boy will leave the balloon on the bus.	1	The kid keeps going around town with the balloon for the rest of the day.	1
The boy tries to board the bus but his balloon keeps following him.	1	He will follow a similar routine to that in clip 1 but the balloon will be able to take care of itself better.	1	The boy is going to let go of the balloon after being told to.	1
He cannot bring the balloon on the bus and he continues to walk.	1	--	--	The bus gate person doesn't allow the boy on, but he leaves the balloon so he can go.	1
--	--	--	--	The boy will go to school again.	1

Table 11 – Predictions and their frequency (*f*) at Probe 7 (Within/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The balloon continues to follow the boy.	10	The boy gets off the bus and reunites with the balloon.	11	The boy gets off the bus and reunites with the balloon.	17
The boy goes to school with the balloon.	7	The balloon continues to follow the boy.	6	The boy arrives at school.	4
The boy gets off the bus and reunites with the balloon.	6	The balloon will have trouble following the bus.	2	The balloon continues to follow the boy.	3
The balloon will have trouble following the bus.	2	The boy gets off the bus.	2	The bus stops.	1
The balloon won't be able to keep up.	1	The boy gets off the bus near his school.	2	The balloon takes a short cut to beat the bus.	1
Boy will get off bus to check out market.	1	The boy will go to school and encounter the creepy man again.	1	Something bad will happen with the balloon.	1
The boy and balloon will continue traveling together, balloon floating on its own.	1	The trip continues and the people are astounded.	1	The balloon stops following the bus.	1
The boy will get off at stop near his home and show all his friends the balloon.	1	The child will arrive at a park.	1	The balloon and boy will reach their destinations but the balloon will start to separate from the boy.	1
The boy gets off the streetcar and the balloon runs away.	1	When he goes to school he will not need to give the balloon to that man to keep.	1	The balloon gets lost.	1
--	--	Someone will pop the balloon.	1	--	--

Table 11 (continued).

--	--	Something will happen to the boy or balloon.	1	--	--
--	--	The boy will now go to school without the balloon and the balloon will meet the by outside when school's over.	1	--	--

Table 12 – Predictions and their frequency (*f*) at Probe 8 (Across/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The man does not catch the balloon.	15	The man catches the balloon.	5	The man does not catch the balloon.	12
The man catches the balloon.	5	The man gives up trying to catch the balloon.	5	The man catches the balloon.	7
The balloon eludes everyone but the boy.	3	The balloon eludes everyone but the boy.	4	The man pops the balloon.	3
The man continues trying to catch the balloon.	2	The man does not catch the balloon.	4	The balloon will find its way into the classroom.	1
The man tries to pop the balloon.	2	Children will try to catch the balloon.	2	People will come to try & catch the balloon but fail. Boy will get balloon after class.	1
The man will become frustrated.	1	The man continues trying to catch the balloon.	2	The balloon leads the man in the hat somewhere.	1
The balloon enters the class.	1	Everyone will want to balloon after school, but it will follow the boy.	1	The balloon would wait for the kids to come out and then let him catch itself.	1
The balloon finds its way to the boy.	1	The balloon will get lost from the boy.	1	The man won't get the balloon. The boy will leave and the balloon will follow him.	1
--	--	The balloon flies away from the school.	1	The boy will leave school and the balloon will stay with him.	1
--	--	The balloon will get inside the facility.	1	The man will try to grab the balloon.	1
--	--	Balloon finds some way to reunite with the boy.	1	The teacher is going to try to contain the balloons & pop it/shove it.	1
--	--	The guy in black grabs and holds onto balloon for the boy.	1	--	--

Table 12 (continued).

--	--	The balloon will protect itself from rain (i.e. he won't need to put it under strangers' umbrellas).	1	--	--
--	--	The balloon escapes yet again.	1	--	--

Table 13 – Predictions and their frequency (*f*) at Probe 9 (Within/Within) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The balloon continues to follow the man until he frees the boy.	8	The balloon continues to follow the man.	12	The balloon continues to follow the man until he frees the boy.	8
The balloon continues to follow the man.	5	The balloon continues to follow the man until he frees the boy.	7	The balloon continues to follow the man.	4
The balloon will get the key to the door.	3	The man catches the balloon.	3	The balloon will get the key to the door.	4
Other people will notice the balloon.	2	Children will try to catch the balloon.	2	The man catches the balloon.	4
The balloon returns to the boy.	2	The balloon will try to get the key and rescue the boy.	1	Headmaster goes to lunch.	1
Children will try to catch the balloon.	2	The man tries to get rid of the balloon.	1	He continues his attempts to catch the balloon.	1
The man catches the balloon.	2	The balloon continues to stalk the headmaster-like guy, but the guy never notices.	1	The mom will get on a bus.	1
The balloon continues to follow the man, but the man is unable to catch it.	1	Man goes into store/shop.	1	The kids try and grab the balloon.	1
Man in the coat turns suddenly and tries to grab the balloon again.	1	Balloon will get revenge on the man.	1	The balloon keeps agitating him, he attacks.	1
The man fetches the boy's mom.	1	He's trying to lose the balloon.	1	The man starts running to get away from the balloon.	1
Balloon will make man unlock the door.	1	--	--	The balloon is now going to annoy the teacher for locking the kid in a room.	1
The balloon will somehow help the boy escape.	1	--	--	The man gets annoyed with the balloon so he goes back and unlocks the kid.	1

Table 13 (continued).

The man will be bothered during his outing & will resort to letting the boy out.	1	--	--	The man goes into a building.	1
--	--	--	--	The balloon will get the man to free the boy.	1

Table 14 – Predictions and their frequency (*f*) at Probe 10 (Within/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The men try to catch the balloon.	7	The men try to catch the balloon.	10	One of the men tries to catch the balloon.	8
The men talk about the balloon.	4	The men talk about the balloon.	4	The men try to catch the balloon.	8
The man will free the boy.	3	The balloon pesters the man.	3	The balloon directs the men back to the boy.	3
The men try to pop the balloon.	3	The men try to pop the balloon.	3	The balloon will get the key to the door.	2
The balloon pesters the men.	2	The balloon reminds the man that the boy is locked away.	2	The men use the boy to get rid of the balloon.	2
The man leaves the other man, and the balloon continues to follow him.	1	The man will free the boy.	2	The school teacher will walk back to the school.	1
The balloon is going to get the key.	1	They ignore the balloon.	1	The balloon goes back and unlocks the door for the boy.	1
The balloon spies on the faculty of the school.	1	The man tries to explain why the balloon follows him all the way.	1	Man plans to get balloon.	1
Man will pretend balloon doesn't exist.	1	The suspicious people will get caught.	1	The balloon will keep following the man.	1
The man will return the balloon to the boy.	1	Balloon finds some way to set the boy free.	1	The man will be confused about the balloon.	1
The older man tries to grab at the balloon.	1	Those two men are involved in some kind of conspiracy (not joking).	1	They are going to talk, but now the topic of the conversation will change to balloon.	1
The balloon will scare the man & make him run away.	1	The two men will come up with a plan to get rid of the annoying balloon.	1	The men find a way to catch the balloon, and maybe destroy it.	1
The two men talk to the boy.	1	--	--	--	--

Table 14 (continued).

The men talking will be distracted by the following balloon.	1	--	--	--	--
They bring the balloon back to the boy to show it follows people.	1	--	--	--	--
The other guy dismisses the principal's experience with the balloon, gets boy out.	1	--	--	--	--

Table 15 – Predictions and their frequency (*f*) at Probe 11 (Across/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy finds the balloon.	13	The boy finds the balloon.	13	The boy finds the balloon.	12
The boy will continue searching for the balloon.	5	The boy chases the balloon.	3	The boy loses the balloon.	5
The boy loses the balloon.	4	The boy loses the balloon.	3	The boy will continue searching for the balloon.	4
The boy and the balloon explore some more.	1	The balloon follows the boy.	2	The balloon looks at mirrors.	2
The balloon is going to hide.	1	The boy will continue searching for the balloon.	2	He explores the market further.	1
The boy wants to buy the painting of the girl.	1	Balloon helps kid get the girl/picture.	1	Someone will attempt to steal the balloon and sell it.	1
The balloon wants to buy a mirror.	1	Boy will find something to buy.	1	I think the kid will leave the market.	1
Someone is going to catch the balloon and the boy will chase them.	1	Boy continues to explore the market.	1	The balloon will find a new owner.	1
The boy will find the balloon & scold it, since he talked to the balloon before.	1	The boy will ditch the balloon for a girl.	1	The boy and the balloon will return home.	1
The balloon continues to fly away and getting the attention of the crowd.	1	Someone else will take the balloon.	1	The balloon will lead the boy into a dangerous situation.	1
The balloon is going to help the little boy find the girl on the image.	1	The kid will be done browsing and be on his way (the balloon too).	1	The boy will chase the balloon.	1

Table 15 (continued).

--	--	The boy will have to start following the balloon.	1	--	--

Table 16 – Predictions and their frequency (*f*) at Probe 12 (Within/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy and girl become friends.	10	The boy and girl become friends.	8	The boy and girl walk together.	10
The boy and girl walk together.	8	The boy and girl walk together.	8	The boy and girl become friends.	7
The boy and girl let go of their balloons.	2	The boy and girl interact with each other.	3	The boy and girl talk with each other.	5
The boy and girl play with each other.	2	The boy and girl spend time with each other.	3	The boy and girl play together.	3
The boy and girl talk with each other.	2	The boy and the girl go their separate ways.	3	The red balloon tries to find other balloons.	2
The boy and the girl will separate and walk away from each other with their balloons.	1	The boy and girl fall in love.	2	The two balloons start to ignore the children.	1
Both the children and balloons are interested in one another.	1	The boy gives the two balloons (Red Blue) back to the girl.	1	The children and the balloons will fall in love.	1
He will give blue balloon back to girl.	1	Gives the blue balloon back to the girl & he continues walking his way.	1	The kids will start going out.	1
Both balloons try to be together.	1	The girl will start to walk away.	1	--	--
The 2 kids start to meet every day.	1	--	--	--	--
The boy and the girl start hanging out together.	1	--	--	--	--

Table 17 – Predictions and their frequency (*f*) at Probe 13 (Across/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy's mother puts the balloon outside.	6	The boys give up and leave.	6	The boys give up and leave.	6
The boys give up and leave.	5	The boy hides from the other boys.	5	The boy releases the balloon outside.	3
The boys try to steal the balloon another way.	3	The girl and the blue balloon arrive.	3	The boys continue trying to steal the balloon.	2
The boy hides from the other boys.	2	The boy closes the window.	2	The boys enter the boy's building.	2
The boy waits until it is safe to go outside again.	2	The boy shares his balloon.	2	The boy's mother puts the balloon outside.	2
The boys continue to wait outside.	2	The boy's mother puts the balloon outside.	2	The group of boys will continue to harass the little boy.	1
He repeats this tomorrow.	1	The boy & balloon will always be friends.	1	The red balloon meets the blue balloon again.	1
The girl will show up.	1	A kid knocks on his door.	1	The kid's parents may still not approve of the balloon.	1
The boy goes to school the next day as well but all the other kids try to steal the balloon.	1	The boys will bully the boy would has balloons affections.	1	The neighbor wants to see the balloon, she knocks on the door.	1
The boy closes the windows.	1	The mob of kids try and take the blue balloon.	1	The group of kids will find the girl with the blue balloon.	1
The girl goes home with her balloon.	1	The boys around aren't going to see the balloon.	1	The boy, and balloon, goes to school the next day.	1
The boy will leave out of a back entrance & find the girl w/ the blue balloon.	1	It cuts to the next day, and the boy walks out with his balloon, but it's even smaller.	1	Once the child sees that the older children has disappeared he will let the balloon go.	1

Table 17 (continued).

The boy closes the doors temporarily then lets the balloon go.	1	Friendship between kid & balloon.	1	This balloon would eventually get stolen.	1
The boy will be more protective about his balloon.	1	The other children in the home will try to take the balloon.	1	Night comes, balloon wanders on its own.	1
The boy keeps the balloon in the house but it tries to escape.	1	Boy goes out again later.	1	The people trying to steal his balloon will soon want to be friends and strike up conversation.	1
Relationship with balloon changes as everyone new leaves.	1	The other boys will beat him up at school.	1	The boy will shut the door.	1
--	--	--	--	The woman scolds him for bringing the balloon inside.	1
--	--	--	--	The kids will fight over to get the balloon.	1
--	--	--	--	He hangs out with his balloon inside.	1
--	--	--	--	The boy will lose the balloon.	1

Table 18 – Predictions and their frequency (*f*) at Probe 14 (Within/Within) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boy buys food.	16	The boy buys food.	16	The boy buys food.	13
The boy tries to buy food.	5	The boy finds that he has no money.	6	The boy finds that he has no money.	6
The boy finds that he has no money.	3	The boy will keep going on more adventures with the balloon.	1	The boy enters the bakery.	3
The boy uses the balloon to get food.	2	The boy will go inside the shop.	1	The boy tries to buy food.	3
The boy walks into the bakery, and the balloon follows him inside.	1	The boy takes one of the snacks.	1	The boy sees how much money he has.	2
The balloon and the boy are going to have lunch.	1	The balloon will cause issues in the bakery too.	1	The boy uses the balloon to get food.	2
The girl will show up.	1	He is checking if he has enough money to buy something.	1	The boy will go into the shop to buy a good and when he comes out his balloon will vanish.	1
The boy goes into the shop to buy something.	1	The boy will eat a pastry. The balloon will leave while this happens.	1	--	--
--	--	The kid doesn't have enough money for the pastries so he trades the balloon.	1	--	--
--	--	Kid won't have enough money, balloon will help.	1	--	--

Table 19 – Predictions and their frequency (*f*) at Probe 15 (Within/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The boys catch up to the boy and the balloon.	8	The boys catch up to the boy and the balloon.	8	The boys catch up to the boy and the balloon.	14
The boy and balloon elude the boys.	5	The boy and balloon elude the boys.	7	The boy and balloon elude the boys.	8
The boy goes home with the balloon.	4	The boys keep searching for the boy and the balloon.	7	The boy and balloon hide from the boys.	3
The boys keep searching for the boy and the balloon.	4	The boy and balloon find the girl with the blue balloon.	3	The boys keep searching for the boy and the balloon.	2
The boy and balloon are helped by an adult.	2	The boys will chase the boy and the balloon until the balloon pops or the boy gets hurt.	1	The older boys will try to get the balloon.	1
The boy and balloon find the girl with the blue balloon.	2	The boy has the man help him hide the balloon.	1	In the fight over the balloon it pops.	1
The gang of boys will pop the balloon.	1	The boy will find a stranger's house that will let him in.	1	The boy will go back to his apartment.	1
The boy will give the balloon to the man of school to keep safely.	1	The boy will share his balloon with a few of the kids.	1	--	--
The man draws suspicion on the children.	1	The boy hides and escapes the crowd of kids, or confronts them in a dead end and the girl with the blue balloon comes and protects the boy and the red balloon.	1	--	--
Boy has to let balloon free to be free from children.	1	--	--	--	--
The boys are going to catch the little boy and destroy his balloon.	1	--	--	--	--

Table 20 – Predictions and their frequency (*f*) at Probe 16 (Across/Across) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The balloon is destroyed.	11	The balloon is destroyed.	13	The balloon is destroyed.	15
The boy becomes upset about losing the balloon.	11	The boy becomes upset about losing the balloon.	4	The boy becomes upset about losing the balloon.	3
The girl with the blue balloon comforts the boy.	2	The boy retrieves the balloon.	4	The boy tries to save the balloon.	3
The owner of the balloon (boy) will run over to the shrinking balloon.	1	The girl with the blue balloon appears.	2	A child who is not the main boy will come pick up the balloon.	1
The boys will run away.	1	The balloon will stay on the ground.	1	The boy leaves the balloon behind and returns home.	1
The boy goes to the balloon, saddened, and tries to "revive" it.	1	The balloon shrivels and the boys turn on the one boy.	1	A dog will save the balloon and the child.	1
The balloon survives and eaves the boy so he can be safe.	1	The balloon then doesn't act like a man. It doesn't play with the boy later.	1	The boy suddenly acknowledges that the balloon is "dead."	1
The balloon will fix its hole by itself.	1	Balloon, dies, kid will take revenge.	1	The child is going to run to it and pick it up.	1
The dog runs in and chases the older boys away.	1	The dog will come and pop the balloon.	1	The girl with the blue balloon would save the boy, because the blue balloon can find the red one.	1
--	--	The balloon "dies" and needs the help of the blue balloon to come back to life.	1	The balloon will lay down on the ground and the boys will disperse.	1
--	--	The kids walk over to it and mourn.	1	The grown up kid comes back and finds the balloon.	1
--	--	--	--	The boy will come take the balloon with him.	1

Table 21 – Predictions and their frequency (*f*) at Probe 17 (Within/Within) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The balloons go to where the boy and his balloon are.	20	The balloons go to where the boy and his balloon are.	21	The balloons go to where the boy and his balloon are.	21
The balloons retaliate against the boys.	6	The balloons retaliate against the boys.	3	The balloons rescue the boy.	3
The balloons are going to save the boy.	1	The balloons continue to float.	2	The balloons retaliate against the boys.	2
Boy sees balloon spirit is not dead, its everywhere.	1	The balloons bind together to revive their fallen comrade.	1	The group of balloons gather over the ground where the mean boys and the kid are.	1
The boy will see all of these balloons and find a new friend.	1	Balloon boy cheers up in a bittersweet way after seeing all those balloons.	1	All the balloons will float up in the sky until they disappear.	1
All other balloons are going to gather together.	1	The balloon survives escape to freedom.	1	The balloons all meet up.	1
--	--	The balloons will mourn for their fallen comrade.	1	The balloons find the boy/dead balloon.	1

Table 22 – Predictions and their frequency (*f*) at Probe 18 (Within/Within) for each level of Orientation.

Situation		Neutral		Fine	
Prediction	*f*	Prediction	*f*	Prediction	*f*
The balloons carry the boy home.	13	The balloons carry the boy home.	7	The balloons carry the boy home.	7
The boy and the balloons continue flying.	6	The boys will see the boy flying with the balloons.	5	The boy and the balloons continue flying.	4
The boys will see the boy flying with the balloons.	3	The boy and the balloons fly around the city.	4	The balloons carry the boy away.	3
The balloons carry the boy to the girl.	2	The boy and the balloons fly to heaven.	2	The balloons take him to the boys.	3
The boy and the balloons fly around the city.	2	The movie ends.	2	The movie ends.	3
The boy lives happily ever after.	2	All the balloons will belong to the boy.	1	The balloons set the boy down on the ground.	2
Picture will fade to black as boy floats away.	1	The kid flies away with the balloons to the school.	1	The boys will see the boy flying with the balloons.	2
The balloons show their appreciation but decide to go back from where they came.	1	The balloons are going to take the boy to the group of kids that bullied him.	1	The boy floats over the mob of kids.	1
--	--	The boy will be happy/move on from his balloon's death.	1	The balloon exploded will "revive."	1
--	--	The boy will be set down gently.	1	The balloons will carry the boy to the road and his balloon will find him.	1
--	--	Boy eventually lands somewhere, taking a balloon with him, but gives it up.	1	The boy will escape his village.	1
--	--	The boy will land and the balloons will return to their owners.	1	Kid could fly to places	1
--	--	Kid happy to be free and away from all evils of his life.	1	The will go to a better place (heaven)?	1

Table 22 (continued).

--	--	The balloons will take him to the girl.	1	--	--
--	--	The boy will find a human friend knowing the balloon will always be with him.	1	--	--

APPENDIX B: ANALYSES OF HIT AND FALSE ALARM RATES

B.1 Hit Rates

Untransformed hit rates were analyzed with a 3 x 3 mixed ANOVA with Probe Type (Within/Within, Within/Across, Across/Across) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Probe Type, $F(2, 162) = 29.53$, $p < .001$, $\eta^2_{partial} = .27$, but no significant effect of Orientation, $F(2, 81) = .60$, $p = .55$, $\eta^2_{partial} = .02$, nor interaction between Probe Type and Orientation, $F(4, 162) = 1.89$, $p = .11$, $\eta^2_{partial} = .05$.

Trend analysis of Probe Type indicated a significant linear, $F(1, 81) = 51.97$, $p < .001$, $\eta^2_{partial} = .39$, but not a significant quadratic trend, $F(1, 81) = 1.53$, $p = .22$, $\eta^2_{partial} = .02$. Thus, participant's hit rates were highest when within a situation and fine event (Within/Within; $M = .85$, $SE = 0.02$), worse when nearing the end of a fine event during an ongoing situation (Within/Across; $M = .75$, $SE = 0.02$), and were lowest when nearing the end of both a fine event and situation (Across/Across; $M = .56$, $SE = 0.03$).

B.2 False Alarm Rates

Untransformed false alarm rates were analyzed with a 3 x 3 mixed ANOVA with Probe Type (Within/Within, Within/Across, Across/Across) as a within-subjects factor and Orientation (Situation, Neutral, Fine) as a between-subjects factor. This analysis revealed a significant main effect of Probe Type, $F(2, 162) = 3.72$, $p = .03$, $\eta^2_{partial} = .04$, but no significant effect of Orientation, $F(2, 81) = 0.82$, $p = .44$, $\eta^2_{partial} = .02$, nor interaction between Probe Type and Orientation, $F(4, 162) = 0.71$, $p = .59$, $\eta^2_{partial} = .02$.

Trend analysis of Probe Type indicated a significant linear, $F(1, 81) = 5.58$, $p = .02$, $\eta^2_{partial} = .06$, but not a significant quadratic trend, $F(1, 81) = 1.19$, $p = .28$, $\eta^2_{partial} = .02$. Thus, participant's false alarm rates were highest when within a situation and fine event (Within/Within; $M = .59$, $SE = 0.03$), lower when nearing the end of a fine event during an ongoing situation (Within/Across; $M = .51$, $SE = 0.03$), and were lowest when nearing the end of both a fine event and situation (Across/Across; $M = .49$, $SE = 0.03$).

APPENDIX C: REPRODUCTION OF MUMMA & DURSO (IN REVISION)

C.1 Experiment 1

The goals of Experiment 1 were 1) to collect and establish the reliability of normative ratings of situations characteristics, which were used to predict judgments of situation boundaries in Experiment 2, and 2) to test the dynamic relationship between changes in situation cues and situation characteristics. The latter is a stronger test of the theoretical relationship between cues and characteristics (i.e., that situation characteristics are interpretations or evaluations of the meaning of a constellation of situation cues), for which evidence has only been provided using static descriptions of situations (Rauthmann et al., 2014). To this end, participants continuously rated a narrative feature film (*The Red Balloon;* Lamorisse, 1956) with respect to one of eight different situation characteristics (i.e., “the Situational Eight DIAMONDS”; Rauthmann et al., 2014). The film was coded by Zacks et al. (2009) for changes in situation cues.

C.1.1 Method

C.1.1.1 Participants

One hundred individuals from the Georgia Institute of Technology participated in partial fulfillment of a research familiarization requirement. Data from one subject was excluded for having no sound during the film and data from two subjects were excluded after each reported that they had seen the film before. The average age of the remaining 97 participants was 20.06 years (*SD* = 2.15), 48 of whom were male. A different set of 12 participants rated seven of the eight DIAMONDS, except for Duty, which had 13 raters.

C.1.1.2 Materials

The stimulus in the present study was the cinematic film, *The Red Balloon* (Lamorisse, 1956), which has been used in previous studies on event segmentation (Kurby, Asiala, & Mills, 2014; Zacks et al., 2009; Zacks et al., 2010). The film is about a young boy who befriends a sentient balloon that attempts to follow the boy throughout his daily activities (e.g., riding the bus, going to school, and going to church). This film was chosen because it includes a variety of changes in situation cues, unfolds continuously in time, has minimal dialogue, and is largely unfamiliar to the participant population. The film was approximately 33 minutes long and was divided into four clips, lasting 7.7, 7.8, 7.4, and 10 minutes, respectively (Zacks et al., 2009).

The Red Balloon was previously coded, frame by frame, by Zacks et al. (2009) for changes in characters (i.e., whenever a different character or characters, including nonhuman animate characters, become the focus of action), interactions between characters (e.g., touching, gesturing, or talking), interactions between characters and objects (e.g., picking up or putting down an object), spatial location (e.g., changes in location or a character's direction of motion), character goals (i.e., whenever a character performed an action associated with a goal that was different from the goal in the previous frame), and causality (i.e., whenever activity in a frame could not be explained by something in a previous frame). Lastly, cuts (i.e., when two continuous film shots abut) were also coded. Correlations among these variables are reported elsewhere (Zacks et al., 2009).

C.1.1.3 Procedure

After providing informed consent, participants were randomly assigned to rate the film with respect to only one of the eight DIAMONDS dimensions (Rauthmann et al., 2014). These characteristics were derived from the factor analysis of ratings from a large and diverse sample of experienced situations (i.e., by people in America and throughout Europe and Asia) and are inclusive of many previous taxonomies of situation characteristics, have good inter-rater reliability, and are tied to a variety of situational cues. The DIAMONDS dimensions comprise Duty (e.g., “Does work need to be done?”), Intellect (e.g., “Is deep cognitive information processing relevant?”), Adversity (e.g., “Is someone under threat?), Mating (e.g., “Is the situation erotically charged?”), pOsitivity (e.g., “Is the situation enjoyable?”), Negativity (e.g., “Could the situation turn negative?”), Deception (e.g., “Is mistrust an issue?”), and Sociality (e.g., “Is meaningful social interaction and relationship building possible?”; Rauthmann & Sherman, 2016).

Participants were seated approximately 61 centimeters away from an 81-centimeter diagonal computer screen. Videos were presented using CARMA software (Girard, 2014), which collects continuous ratings of audiovisual stimuli at 1-second intervals. During the experiment, participants indicated the extent to which they felt the current situation contained one of the DIAMONDS dimensions: either Duty (i.e., “Work, tasks, or duties”), Intellect (i.e., “Intellectual, aesthetic, profound things”), Adversity (i.e., “Threat, accusation, criticism”), Mating (i.e., “Romance, sexuality, love”), pOsitivity (i.e., “Positive, pleasant, nice things”), Negativity (i.e., “Negative things, unpleasant things, bad feelings”), Deception (i.e., “Deceit, lie, dishonesty”), or Sociality (i.e., “Communication, interaction, social relationships”). Participants were told that “The ‘current situation’ is whatever you think the current circumstances or state of affairs in the film is.”

Participants used a joystick to indicate their response on a 7-point scale, which ranged from “Not at All” (1) to “Extremely” (7). The 7-point scale and characteristic-nouns came from the S8-II (Rauthmann & Sherman, 2016), which is a validated brief measure of the DIAMONDS dimensions. Participants rated the film by moving the joystick away from them as the dimension of interest increased (i.e., towards “Extremely”) and toward them as it diminished (i.e., towards “Not at All”). The rating scale and the participant’s current position on the scale were displayed on the screen adjacent to the film. Participants could refer to a printed copy of the nouns corresponding to their assigned characteristic throughout the experimental task. After rating each clip, participants took a brief break before continuing to the next of the four clips, before which task instructions were repeated.

Before rating The Red Balloon, a 4.4-minute-long clip from North by Northwest (Hitchcock, 1959) was used for practice, during which participants were instructed to rate the extent to which the current situation contains “Interesting, stimulating, or engaging things.”

C.1.2 Results and Discussion

C.1.2.1 Ratings of Characteristics

Inter-rater reliability of the ratings of each situation characteristic was assessed with intraclass correlations. For each participant, ratings were averaged over every five contiguous 1-second time bins. 5-second time bins were used to be consistent with previous analyses of segmentation behavior while viewing *The Red Balloon* (Kurby et al., 2014; Zacks et al., 2009; Zacks et al., 2010). Because the present study was interested in how changes in characteristics are perceived normatively (i.e., on average), a two-way random-average-measures consistency model was obtained for each characteristic (ICC[C,k]; McGraw & Wong, 1996). Table 23 provides the intraclass correlation coefficients and 95% confidence intervals for each of the eight DIAMONDS dimensions, all of which correspond to excellent reliability (Cicchetti, 1994). Thus, consistency was high among participants continuously rating the same situation characteristic throughout the film.

Table 23 – Intraclass correlation coefficients and 95% confidence intervals for continuous ratings of the eight DIAMONDS dimensions.

DIAMONDS	n	Intraclass Correlation Coefficient	95% CI
Duty	13	.80***	[.77, .83]
Intellect	12	.92***	[.91, .94]
Adversity	12	.91***	[.90, .92]
Mating	12	.83***	[.81, .86]
pOsitivity	12	.94***	[.93, .95]
Negativity	12	.90***	[.88, .91]
Deception	12	.78***	[.74, .81]
Sociality	12	.91***	[.90, .92]

Note. n = number of raters. CI = confidence interval.

*** $p < .001$

C.1.2.2 Characteristic-Change

To measure the momentary change in each characteristic throughout the film, the absolute value of the mathematical derivative of the rating value was approximated for each 1-second time bin, then averaged over each successive five bins. Thus, an average derivative of zero indicates that no change in a characteristic occurred in those 5 seconds, whereas bins containing a non-zero average derivative indicate the extent to which change in a characteristic did occur. For each situation characteristic, the mean derivative for each subject (averaged over all time bins) was computed and collectively compared to zero using a one-sample *t*-test. Table 24 indicates the grand mean of the derivative for each situation characteristic, all of which were significantly greater than zero, indicating that no characteristic remained unchanged throughout the film

Table 24 – Average derivative and standard error of DIAMONDS dimensions.

DIAMONDS	Mean Derivative	*SE*
Duty	0.11**	0.02
Intellect	0.08***	0.01
Adversity	0.05***	0.01
Mating	0.04**	0.01
pOsitivity	0.08***	0.01
Negativity	0.06***	0.01
Deception	0.09***	0.02
Sociality	0.15***	0.01

****** $p < .01$, *** $p < .001$

C.1.2.3 Predicting Changes in Characteristics

A major goal of Experiment 1 was to determine whether changes in situation cues were generally predictive of changes in characteristics. In the present studies, situation cues refer to characters (Who?), their interactions with other characters and objects (What?), their motivations (i.e., causes of activity, such as goals; Why?), and space (Where?). To predict changes in characteristics (i.e., the eight DIAMONDS dimensions) from cuts and changes in these cues, eight separate linear mixed models were built using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in R statistical software (R Core Team, 2013). The statistical significance of fixed effects was obtained using the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2015), which uses Satterthwaite's approximation for degrees of freedom. Each model included seven fixed effects (i.e., one binary predictor variable coding for the presence of a cut in each time bin and one binary predictor for each of the six cues coding for a change in that cue) and a random intercept for participants to account for the fact that repeated observations were made on each participant.

Table 25 shows the coefficients for the seven predictor variables for each of the eight DIAMONDS dimensions. Indeed, changes in all eight DIAMONDS dimensions were positively related to changes in at least one of the six situation cues. In particular, participants perceived changes in all of the DIAMONDS dimensions when character(s) changed. That changes in characters were predictive of changes in all eight characteristics is consistent with the idea that characters are fundamental to situation model construction

(Bailey et al., 2017; Radvansky & Dijkstra, 2007; Thierrault, Rinck, & Zwaan, 2006; Zwaan & Radvansky, 1998).

Table 25 – Fixed effects coefficients and significance of cuts and changes in cues in predicting changes in characteristics.

		Situation Cues					
DIAMONDS	Cuts	Character	Character with Character[a]	Character with Object[b]	Cause	Goal	Space
Duty	0.009	0.029***	0.020**	0.013	0.007	-0.013	0.006
Intellect	-0.008	0.019***	-0.003	0.005	0.012*	-0.005	0.007
Adversity	-0.006	0.026***	0.010*	0.004	0.008	-0.007	0.002
Mating	0.001	0.008*	-0.003	0.004	0.008*	0.007	0.002
pOsitivity	-0.003	0.025***	0.013*	0.021***	0.012*	0.006	-0.003
Negativity	-0.015***	0.034***	0.009	0.008	0.006	-0.004	0.010*
Deception	-0.004	0.033***	0.010	0.010	0.004	-0.013	0.018**
Sociality	-0.010	0.094***	0.038***	0.023*	0.004	-0.011	0.009

[a]"Character with Character" means changes in physical interactions between characters.
[b]"Character with Object" means physical interactions between characters and objects.
* $p < .05$, ** $p < .01$, *** $p < .001$

Apart from changes in characters, changes in other situation cues occasionally predicted changes in multiple characteristics; for example, changes in the interactions between characters were positively related to changes in Duty, Adversity, pOsitivity, and

Sociality. In short, situation characteristics are at least partly associated with situation cues, as many have theorized (Endler, 1981; Rauthmann et al., 2015) but have not demonstrated with dynamic stimuli.

C.2 Experiment 2

Having found support for the hypothesis that situation characteristics are dynamically tied to a variety of situation cues in Experiment 1, the goal of Experiment 2 was to test two competing perspectives on how changes in cues and characteristics might correspond to the perception of a new situation beginning – the objectivist and subjectivist perspectives.

The objectivist perspective emphasizes the role of the external features of situations in influencing behavior. For example, much of experimental social psychology considers manipulation of situation cues sufficient to create different situations (e.g., by varying who is present, such as the number of bystanders; Horstmann, Rauthmann, & Sherman, 2017). Indeed, judgments about when situations change during narrative stories, film, or video games correspond to changes in characters (“Who?”; Magliano, Kopp, McNerny, Radvansky, & Zacks, 2012; Magliano, Radvansky, Forsythe, & Copeland, 2014), their goals (“Why?”; Magliano et al., 2012; Magliano et al., 2014; Magliano, Taylor, & Kim, 2005), spatial-temporal framework (“When?/Where?”; Magliano et al., 2001; Magliano et al., 2011; Magliano et al., 2014), and events (“What?”; Magliano et al., 2012). Thus, the objectivist tradition would predict that changes in situation cues will leave little variance in situation-change judgments to be accounted for by changes in situation characteristics (e.g., the situation's positivity or adversity).

The subjectivist perspective emphasizes the role of the phenomenological aspects of situations in influencing behavior, arguing that people ultimately react to the characteristics or qualities of situations, not their situation cues per se (Edwards & Templeton, 2005; Endler, 1981). For example, recent taxonomies of situation characteristics (Parrigon et al., 2017; Sherman, Rauthmann, Brown, Serfass, & Jones, 2015; Rauthmann et al., 2014) predict the behaviors of people in situations from the situation's characteristics in conceptually meaningful ways (e.g., situations perceived to be "important" tend to elicit conscientious behavior; Parrigon et al. 2017). Thus, the subjectivist perspective would predict that changes in situation characteristics should leave little variance in situation-change judgments to be accounted for by changes in situation cues.

Lastly, if changes in situation characteristics are at all predictive of situation-change judgments, then situation model updating effects should be observed. Specifically, The Event Indexing Model (Zwaan & Radvansky, 1998) predicts that the more a situation changes, the greater the probability of perceiving that a new situation has begun (i.e., the "additivity hypothesis"; Magliano et al., 2001; Rinck & Weber, 2003; Zacks et al., 2009; Zwaan et al., 1995; Zwaan et al., 1998; Zwaan & Radvansky, 1998). Thus, the additivity hypothesis predicts that the more situation characteristics are changing at any given moment in time, the greater the probability that one will perceive that a new situation is beginning.

C.2.1 Method

C.2.1.1 Participants

Fifty-seven additional participants from the Georgia Institute of Technology participated in partial fulfillment of a research familiarization requirement or for pay. Data for one participant was excluded because of an audio malfunction throughout the film. The remaining 56 participants (25 male) had a mean age of 20.45 years (SD = 2.41).

C.2.1.2 Procedure

After providing informed consent, participants were seated approximately 2 feet away from a 32" inch diagonal computer screen. Stimuli were presented with PsychoPy (Pierce, 2007). To measure the perception of situation-change, a modified version of an event-marking paradigm was used (Newtson, 1973; Zacks, Tversky, & Iyer, 2001). Typically, participants are instructed to parse activity into either the smallest or largest units that are natural and meaningful to them (i.e., fine and coarse segmentation, respectively; Newtson, 1973; Zacks et al., 2001). In the present study, participants were instructed to press the spacebar every time they felt that a new situation had begun. The same practice clip from Experiment 1 was used for practicing the segmentation task in Experiment 2. The precise moment that the spacebar was pressed each time during the film was recorded and was later assigned to the 1-second time bin in which it occurred.

C.2.2 Results and Discussion

C.2.2.1 Situation Boundaries

The mean number of situation boundaries identified during *The Red Balloon* was 75.79 (*SE* = 6.06), with a median of 62. To measure the extent of agreement on the boundaries between situations, a "segmentation norm" was created by calculating the proportion of participants who identified a situation boundary in each time bin (Figure 1). Next, a point-biserial correlation was computed between each participant's (binary) segmentation data and the segmentation norm (Kurby & Zacks, 2011). The resulting correlation coefficients ($\bar{r}$ = .381, *SE* = 0.02) were significantly larger than zero, *t*(55) = 24.16, p < .001.

C.2.2.2 Predicting Situation Boundaries

To predict situation boundary judgments, a logistic mixed-effects model was created using the *lme4* package (Bates et al., 2015) in R statistical software. A full model was constructed, which comprised a binary predictor for cuts, six binary predictors for changes in each of the six situation cues, eight continuous predictors for changes in each of the eight situation characteristics, and a random intercept for participants to account for the fact that repeated observations were made on each participant. A normative measure of how much each of the eight DIAMONDS dimensions were changing in each time bin was computed as follows: the absolute value of the derivative of a participant's rating of a characteristic was calculated for each 1-second time bin and averaged over each successive five time bins, and then that value was averaged with those of other participants rating that characteristic for each of the 5-second time bins. All continuous predictors were standardized prior to analysis.

Regarding the competing hypotheses of the objectivist and subjectivist perspectives, log-likelihood ratio tests were performed to determine whether changes in cues could account for additional variance beyond that accounted for by changes in characteristics (and vice versa). To this end, two reduced models were created from the full model described previously - one without predictors for cues and one without predictors for characteristics. Regarding the contributions of changes in cues and characteristics in predicting situation boundary judgments, both reduced models fit the data significantly worse than the full model (Table 26), suggesting that changes in situation cues and situation characteristics account for unique variance in situation boundary judgments.

Table 26 – Logistic mixed-effect models for situation boundary judgments and log-likelihood ratio tests between full and reduced models.

Model	Equation	Log-Likelihood	χ^2 (df)	*p*
Full	Situation Change = Cuts + Cues + Characteristics	-9353.7	–	–
Reduced[a]	Situation Change = Cuts + Cues	-9513.6	319.63 (8)	< .001***
Reduced[b]	Situation Change = Cuts + Characteristics	-9633.5	559.52 (6)	< .001***

[a]When compared to full model, assesses contribution of changes in characteristics to model fit.
[b]When compared to full model, assesses contribution of changes in cues to model fit.
*** $p < .001$

In the full model (Table 27), the likelihood of identifying a situation boundary increased with changes in cues (i.e., characters, their interactions with other characters or objects, causality, and space) and characteristics (i.e., Duty, Intellect, pOsitivity, Negativity, Deception, and Sociality). It is notable that changes in Duty, pOsitivity, Negativity, and Sociality were predictive of situation boundaries as these characteristics are among the most prominent and replicable in the literature on situational taxonomies (Rauthmann, 2015). Separately, the presence of a cut also increased the likelihood of perceiving a situation boundary, which differs from previous analyses of *The Red Balloon* (Kurby et al., 2014; Zacks et al., 2009; Zacks et al., 2010), where cuts were mostly negatively associated with the identification of fine and coarse-grained event boundaries. Even in the reduced model without situation characteristics, cuts were positively associated with (Odds ratio = 1.10; $z = 2.17$, $p = .03$). Thus, cuts may have a different effect on the perception of situation boundaries than on the perception of event boundaries (i.e., the smallest or largest meaningful units of activity). In short, the results of Experiment 2 suggest a compromise between the objectivist and subjectivist perspectives – changes in situation cues and characteristics are both important for predicting situation-change judgments and do so independently.

Table 27 – Odds ratios and significance of cuts, cue-changes, and characteristic-changes in predicting situation boundaries.

Predictors	Odds Ratio	*z*	*p*
Cuts	1.12	2.40	0.017*
Situation Cues			
Character	1.53	9.29	< 0.001***
Character with Character[a]	1.45	7.44	< 0.001***
Character with Object[b]	1.23	3.77	< 0.001***
Cause	1.50	8.56	< 0.001***
Goal	1.05	0.95	0.342
Space	1.14	2.86	0.004**
DIAMONDS			
Duty	1.13	6.45	< 0.001***
Intellect	1.12	5.36	< 0.001***
Adversity	1.01	0.65	0.517
Mating	1.02	1.00	0.319
pOsitivity	1.10	4.10	< 0.001***
Negativity	1.15	5.90	< 0.001***
Deception	1.04	2.02	0.043*
Sociality	1.08	3.62	< 0.001***

[a]"Character with Character" means changes in physical interactions between characters.
[b]"Character with Object" means physical interactions between characters and object.
* $p < .05$, ** $p < .01$, *** $p < .001$

C.2.2.3 Additivity Hypothesis

That changes in situation characteristics predict situation boundaries independently of changes in situation cues suggests that characteristics constitute meaningful facets of situation models. To test the additivity hypothesis (i.e., the more that characteristics are changing in any given moment, the greater the probability of perceiving that a new situation is beginning), two new predictor variables were created that summed the number of changes in all six situation cues and average derivatives in all situation characteristics in each time bin, respectively. All continuous predictors were mean-centered prior to analysis. Cuts (Odds ratio = 1.20, $z = 4.71$, $p < .001$), total number of changes in cues (Odds ratio = 1.36, $z = 23.89$, $p < .001$), and total amount of change in characteristics (Odds ratio = 2.23, $z = 18.50$, $p < .001$) each increased the likelihood of identifying a situation boundary. Thus, not only did the present study replicate the additive effects of changes in situation cues on the probability of identifying a situation boundary, it also suggests that the additive effect of changes in characteristics manifest independently of the additive effect of changes in cues.

C.2.2.4 Fine and Coarse Segmentation

We determined whether changes in situation cues and characteristics are also related to event boundaries obtained under more traditional segmentation instructions (i.e., segmenting an activity into either the largest or smallest units that are natural and meaningful; Newtson, 1973). To this end, we obtained and predicted fine and coarse-grained segmentation from a previous study using *The Red Balloon* (Kurby et al., 2014) using exactly the same statistical models from Experiment 2. The results of these analyses (Tables 28 and 29 for fine- and coarse-event boundaries, respectively) collectively suggest that changes in situation cues and characteristics independently increase the likelihood of perceiving any kind event boundary (i.e., coarse, situation, or fine).

Table 28 – Odds ratios and significance of cuts, cue-changes, and characteristic-changes in predicting fine-event boundaries.

Predictors	Odds Ratio	*z*	*p*
Cuts	0.98	-0.38	0.702
Situation Cues			
Character	1.22	4.12	< 0.001***
Character with Character[a]	1.16	2.66	0.008**
Character with Object[b]	1.39	5.59	< 0.001***
Cause	1.02	0.31	0.759
Goal	1.06	1.06	0.290
Space	1.37	6.61	< 0.001***
DIAMONDS			
Duty	1.12	5.37	< 0.001***
Intellect	1.06	2.75	0.006**
Adversity	1.09	3.43	0.006**
Mating	1.11	4.65	< 0.001***
pOsitivity	1.06	2.32	0.021*
Negativity	1.09	3.24	< 0.001***
Deception	1.03	1.36	0.175
Sociality	1.08	3.21	< 0.001***

[a]"Character with Character" means changes in physical interactions between characters.
[b]"Character with Object" means physical interactions between characters and object.
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 29 – Odds ratios, significance of cuts, cue-changes, and characteristic-changes in predicting coarse-event boundaries.

Predictors	Odds Ratio	*z*	*p*
Cuts	0.84	-2.59	0.010*
Situation Cues			
Character	1.54	6.43	< 0.001***
Character with Character[a]	1.35	4.17	< 0.001***
Character with Object[b]	1.35	5.02	< 0.001***
Cause	1.47	3.47	< 0.001***
Goal	1.27	2.60	0.009**
Space	1.39	4.69	< 0.001***
DIAMONDS			
Duty	1.14	4.83	< 0.001***
Intellect	1.21	6.79	< 0.001***
Adversity	1.08	2.35	0.019*
Mating	1.03	1.03	0.301
pOsitivity	1.07	1.99	0.047*
Negativity	1.10	2.94	0.003**
Deception	0.99	-0.21	0.834
Sociality	1.06	1.80	0.072

[a]"Character with Character" means changes in physical interactions between characters.
[b]"Character with Object" means physical interactions between characters and object.
* $p < .05$, ** $p < .01$, *** $p < .001$

REFERENCES

Abbott, V., Black, J. B., & Smith, E. E. (1985). The representation of scripts in memory. *Journal of Memory and Language*, *24*(2), 179-199.

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors*, *37*(1), 85-104.

Allison, P. D. (1999). *Logistic regression using SAS: Theory and application*. Cary, NC: SAS Institute.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829-839.

Bailey, H. R., Kurby, C. A., Giovannetti, T., & Zacks, J. M. (2013). Action perception predicts action performance. *Neuropsychologia*, *51*(11), 2294-2304.

Bailey, H. R., Kurby, C. A., Sargent, J. Q., & Zacks, J. M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. *Memory & Cognition*, *45*(6), 940-955.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, *95*(3), 709-721.

Bartlett, FC (1932). Remembering: A study in experimental and social psychology. New York: The Macmillan Company.

Battistich, V. A., & Thompson, E. G. (1980). Students' perceptions of the college milieu: A multidimensional scaling analysis. *Personality and Social Psychology Bulletin*, *6*(1), 74-82.

Bedny, G. and Meister, D., 1999, Theory of activity and situation awareness. *International Journal of Cognitive Ergonomics*, *3*(1), pp. 63-72.

Bläsing, B. E. (2015). Segmentation of dance movement: Effects of expertise, visual familiarity, motor experience and music. *Frontiers in Psychology, 5*, 1500.

Bohn-Gettler, C. M. (2014). Does monitoring event changes improve comprehension? *Discourse Processes, 51*(5-6), 398-425.

Bower, G. H., & Clark-Meyers, G. (1980). Memory for scripts with organized vs. randomized presentations. *British Journal of Psychology, 71*(3), 369-377.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology, 11*(2), 177-220.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology, 3*(2), 193-209.

Brewer, W. F., & Dupree, D. A. (1983). Use of plan schemata in the recall and recognition of goal-directed actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(1), 117-129.

Brown, G. S., & White, K. G. (2005). The optimal correction for estimating extreme discriminability. *Behavior Research Methods, 37*(3), 436-449. Cambridge: Cambridge University Press.

Christoffersen, K., Woods, D. D., & Blike, G. T. (2007). Discovering the events expert practitioners extract from dynamic data streams: The modified unit marking technique. *Cognition, Technology & Work, 9*(2), 81-98.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284-290.

Cohen, C. E., & Ebbesen, E. B. (1979). Observational goals and schema activation: A theoretical framework for behavior perception. *Journal of Experimental Social Psychology, 15*(4), 305-329.

Cook, A. E., Limber, J. E., & O'Brien, E. J. (2001). Situation-based context and the availability of predictive inferences. *Journal of Memory and Language, 44*(2), 220-234.

Corwin, J. (1994). On measuring discrimination and response bias: Unequal numbers of targets and distractors and two classes of distractors. *Neuropsychology, 8*(1), 110 117.

Dekker, S., & Lützhöft, M. (2004). Correspondence, cognition and sensemaking: A radical empiricist view of situation awareness. In S. Banbury & S. Tremblay (Eds.), A cognitive approach to situation awareness: Theory, measurement and application. Aldershot, UK: Ashgate Publishing.

Delogu, F., Drenhaus, H., & Crocker, M. W. (2018). On the predictability of event boundaries in discourse: An ERP investigation. *Memory & Cognition, 46*(2), 315-325.

Deutsch, S. E., Pew, R. W., Rogers, W. H., & Tenney, Y. J. (1994). Toward a Methodology for Defining Situation Awareness Requirements—A Progress Report. National Aeronautics and Space Administration, BBN Report, (7983).

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2008). Time travel through language: Temporal shifts rapidly decrease information accessibility during reading. *Psychonomic Bulletin & Review*, *15*(4), 750-756.

Doane, S. M., Sohn, Y. W., & Jodlowski, M. T. (2004). Pilot ability to anticipate the consequences of flight actions as a function of expertise. *Human factors*, *46*(1), 92-103.

Durso, F. T., & Alexander, A. L. (2010). Managing workload, performance, and situation awareness in aviation systems. In Human factors in aviation (pp. 217-247). Academic Press.

Durso, F. T., & Gronlund, S. D. (1999). Situation awareness. Handbook of applied cognition, 283-314.

Durso, F. T., & Sethumadhavan, A. (2008). Situation awareness: Understanding dynamic environments. *Human Factors*, *50*(3), 442-448.

Durso, F. T., Bleckley, M. K., & Dattel, A. R. (2006). Does situation awareness add to the validity of cognitive tests? *Human Factors*, *48*(4), 721-733.

Durso, F. T., Dattel, A. R., & Pop, V. L. (2018). 20 Expertise and Transportation. The Cambridge Handbook of Expertise and Expert Performance, 356.

Durso, F. T., Johnson, B. R., & Crutchfield, J. M. (2010). Dimensions of air traffic control tower information needs: From information requests to display design. *Journal of Experimental Psychology: Applied*, *16*(3), 219-237.

Durso, F. T., Rawson, K. A., & Girotto, S. (2007). Comprehension and situation awareness. In F. T. Durso, R. S. Nickerson, S. T. Dumais, S. Lewandowsky, & T. J. Perfect (Eds.), Handbook of applied cognition (2nd ed., pp. 163–193). Chichester, UK: Wiley.

Durso, F. T., Truitt, T. R., Hackworth, C. A., Crutchfield, J. M., Nikolic, D., Moertl, P. M., Ohrt, D. D., & Manning, C. A. (1995). Expertise and chess: A pilot study comparing situation awareness methodologies. In D. J. Garland & M. R. Endsley (Eds.), Experimental analysis and measurement of situation awareness (pp. 295–303). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Eckes, T. (1995). Features of situations: A two-mode clustering study of situation prototypes. *Personality and Social Psychology Bulletin, 21*(4), 366-374.

Edwards, J. A., & Templeton, A. (2005). The structure of perceived qualities of situations. *European Journal of Social Psychology, 35*(6), 705-723.

Eisenberg, M. L., & Zacks, J. M. (2016). Ambient and focal visual processing of naturalistic activity. *Journal of Vision, 16*(2):5, 1-12.

Eisenberg, M. L., Zacks, J. M., & Flores, S. (2018). Dynamic prediction during perception of everyday events. *Cognitive Research: Principles and Implications, 3*(1), 53.

Endsley, M. R. (1990). Situation awareness in dynamic human decision making: Theory and measurement (Doctoral dissertation, University of Southern California).

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R. (2000, July). Situation models: An avenue to the modeling of mental models. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, No. 1, pp. 61-64). Sage CA: Los Angeles, CA: SAGE Publications.

Engström, J., Bärgman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G. B., & Victor, T. (2018). Great expectations: A predictive processing account of automobile driving. *Theoretical Issues in Ergonomics Science, 19*(2), 156-194.

Flach, J. M. (1995). Situation awareness: Proceed with caution. *Human Factors, 37*(1), 149-157.

Flach, J. M. (1996). Situation awareness: In search of meaning. *CSERIAC Gateway, 6*(6), 1–4.

Flach, J. M., Feufel, M. A., Reynolds, P. L., Parker, S. H., & Kellogg, K. M. (2017). Decision making in practice: The dynamics of muddling through. *Applied Ergonomics, 63*, 133-141.

Flach, J., Mulder, M. & Paassen, M.V., (2004). *The concept of the situation in psychology. In A cognitive approach to situation awareness: Theory and application.* In S. Banbury and S. Tremblay (Eds.), pp. 42–60 (Aldershot, UK: Ashgate).

Flores, S., Bailey, H. R., Eisenberg, M. L., & Zacks, J. M. (2017). Event segmentation improves event memory up to one month later. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(8), 1183-1202.

Forgas, J. P. (1976). An unobtrusive study of reactions to national stereotypes in four European countries. *The Journal of Social Psychology*, *99*(1), 37-42.

Forgas, J. P. (1982). Episode cognition: Internal representations of interaction routines. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 15, pp. 59-100). San Diego, CA: Academic Press.

Foroughi, C. K., Werner, N. E., Barragán, D., & Boehm-Davis, D. A. (2015). Interruptions disrupt reading comprehension. *Journal of Experimental Psychology: General*, *144*(3), 704-709.

Furr, R. M., & Funder, D. C. (2004). Situational similarity and behavioral consistency: Subjective, objective, variable-centered, and person-centered approaches. *Journal of Research in Personality, 38*(5), 421-447.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*(1), 69-83.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371-395.

Graziano, W. G., Moore, J. S., & Collins, J. E. (1988). Social cognition as segmentation of the stream of behavior. *Developmental Psychology, 24*(4), 568-573.

Gronlund, S. D., & Neuschatz, J. S. (2014). Eyewitness identification discriminability: ROC analysis versus logistic regression. *Journal of Applied Research in Memory and Cognition, 3*(2), 54-57.

Hanson, C., & Hirst, W. (1989). On the representation of events: A study of orientation, recall, and recognition. *Journal of Experimental Psychology: General, 118*(2), 136-147.

Hard, B. M., Meyer, M., & Baldwin, D. (2019). Attention reorganizes as structure is detected in dynamic action. *Memory & Cognition, 47*(1), 17-32.

Hard, B. M., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of Experimental Psychology: General, 140*(4), 586-604.

Hard, B. M., Tversky, B., & Lang, D. S. (2006). Making sense of abstract events: Building event schemas. *Memory & Cognition, 34*(6), 1221-1235.

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers, 27*(1), 46-51.

Hitchcock, A. (Director) (1959). North by northwest [Motion picture]. United States: Metro Goldwyn-Mayer.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*(5), 849-878.

Huff, M., Meitz, T. G., & Papenmeier, F. (2014). Changes in situation models modulate processes of event perception in audiovisual narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(5), 1377-1388.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*(2), 386-408.

Kadlec, H. (1999). Statistical properties of d' and β estimates of signal detection theory. *Psychological Methods, 4*(1), 22-43.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review, 95*(2), 163-182.

Kosie, J. E., & Baldwin, D. (2019). Attentional profiles linked to event segmentation are robust to missing information. *Cognitive research: Principles and Implications, 4*(1), 8.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences, 12*(2), 72-79.

Kurby, C. A., & Zacks, J. M. (2011). Age differences in the perception of hierarchical structure in events. *Memory & Cognition, 39*(1), 75-91.

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition, 40*(5), 812-826.

Kurby, C. A., & Zacks, J. M. (2018). Preserved neural event segmentation in healthy older adults. *Psychology and Aging*, *33*(2), 232-245.

Kurby, C. A., & Zacks, J. M. (2019). Age differences in the perception of goal structure in everyday activity. *Psychology and Aging, 34*(2), 187-201.

Kurby, C. A., Asiala, L. K., & Mills, S. R. (2014). Aging and the segmentation of narrative film. *Aging, Neuropsychology, and Cognition, 21*(4), 444-463.

Lamorisse, A. (Director) (1956). The red balloon [Motion picture]. France: Films Montsouris.

Landman, A., Groen, E. L., Van Paassen, M. M., Bronkhorst, A. W., & Mulder, M. (2017). Dealing with unexpected events on the flight deck: A conceptual model of startle and surprise. *Human Factors*, *59*(8), 1161-1172.

Lassiter, G. D., Geers, A. L., Apple, K. J., & Beers, M. J. (2000). Observational goals and behavior unitization: A reexamination. *Journal of Experimental Social Psychology*, *36*(6), 649-659.

Levin, D. T., & Baker, L. J. (2017). Bridging views in cinema: A review of the art and science of view integration. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(5), e1436.

Levine, D., Hirsh-Pasek, K., Pace, A., & Michnick Golinkoff, R. (2017). A goal bias in action: The boundaries adults perceive in events align with sites of actor intent. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(6), 916-927.

Lichtenstein, E. H., & Brewer, W. F. (1980). Memory for goal-directed events. *Cognitive Psychology, 12*(3), 412-445.

Lorch, R. F., & Myers, J. L. (1990). Regression analyses of repeated measures data in Cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(1), 149-157.

Ma, R., & Kaber, D. B. (2007). Situation awareness and driving performance in a simulated navigation task. *Ergonomics, 50*(8), 1351-1364.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and "nonparametric" indexes. *Psychological Bulletin, 107*(3), 401-413.

Macmillan, N. A., & Creelman, C. D. (1996). Triangles in ROC space: History and theory of “nonparametric” measures of sensitivity and response bias. *Psychonomic Bulletin & Review, 3*(2), 164-170.

Macmillan, N. A., Rotello, C. M., & Miller, J. O. (2004). The sampling distributions of Gaussian ROC statistics. *Perception & Psychophysics, 66*(3), 406-421.

Magliano, J. P., & Graesser, A. C. (1991). A three-pronged method for studying Inference generation in literary text. *Poetics, 20*(3), 193-232.

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences While viewing a movie. *Discourse Processes, 22*(3), 199-224.

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 15*(5), 533-545.

Magliano, J. P., Radvansky, G. A., Forsythe, J. C., & Copeland, D. E. (2014). Event segmentation during first-person continuous events. *Journal of Cognitive Psychology, 26*(6), 649-661.

Magliano, J. P., Skowronski, J. J., Britt, M. A., Güss, C. D., & Forsythe, C. (2008). What do you want? How perceivers use cues to make goal inferences about others. *Cognition, 106*(2), 594-632.

Magliano, J. P., Taylor, H. A., & Kim, H. J. J. (2005). When goals collide: Monitoring the goals of multiple characters. *Memory & Cognition, 33*(8), 1357-1367.

Magliano, J., Kopp, K., McNerney, M. W., Radvansky, G. A., & Zacks, J. M. (2012). Aging and perceived event structure as a function of modality. *Aging, Neuropsychology, and Cognition, 19*(1-2), 264-282.

Magnusson, D. (1971). An analysis of situational dimensions. *Perceptual and Motor Skills, 32*(3), 851-867.

Markus, H., Smith, J., & Moreland, R. L. (1985). Role of the self-concept in the perception of others. *Journal of Personality and Social Psychology, 49*(6), 1494-1512.

Massad, C. M., Hubbard, M., & Newtson, D. (1979). Selective perception of events. *Journal of Experimental Social Psychology, 15*(6), 513-532.

McDaniel, M. A., Schmalhofer, F., & Keefe, D. E. (2001). What is minimal about predictive inferences? *Psychonomic Bulletin & Review, 8*(4), 840-846.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation, 51*, 297-384.

Migueles, M., & García-Bajos, E. (2012). The power of script knowledge and selective retrieval in the recall of daily activities. *The Journal of General Psychology, 139*(2), 100-113.

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language, 26*(2), 165-187.

Newberry, K. M., & Bailey, H. R. (2019). Does semantic knowledge influence event segmentation and recall of text? *Memory & Cognition*, 1-15.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, *28*(1), 28-38.

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology, 12*(5), 436-450.

Newtson, D., Engquist, G. A., & Bois, J. (1977). The objective basis of behavior units. *Journal of Personality and Social Psychology, 35*(12), 847-862.

Niessen, C., Eyferth, K., & Bierwagen, T. (1999). Modelling cognitive processes of experienced air traffic controllers. *Ergonomics, 42*(11), 1507-1520.

Niesser, U., (1976). *Cognition and Reality: Principles and Implications of Cognitive Psychology* (San Francisco: Freeman).

O'Brien, K. S., & O'Hare, D. (2007). Situational awareness ability and cognitive skills training in a complex real-world task. *Ergonomics*, *50*(7), 1064-1091.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140-160.

Parrigon, S., Woo, S. E., Tay, L., & Wang, T. (2017). CAPTION-ing the situation: A lexically derived taxonomy of psychological situation characteristics. *Journal of Personality and Social Psychology, 112*(4), 642-681.

Pastore, R. E., Crawley, E. J., Berens, M. S., & Skelly, M. A. (2003). “Nonparametric” A’ and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review, 10*(3), 556-569.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8-13.

Peracchi, K. A., & O’Brien, E. J. (2004). Character profiles and the activation of predictive inferences. *Memory & Cognition, 32*(7), 1044-1052.

Pervin, L. A. (1978). Definitions, measurements, and classifications of stimuli, situations, and environments. *Human Ecology*, *6*(1), 71-105.

Pollack, I., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science, 1*(1-12), 125-126.

Radvansky, G. A. (2005). Situation models, propositions, and the fan effect. *Psychonomic Bulletin & Review*, *12*(3), 478-483.

Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science*, *21*(4), 269-272.

Radvansky, G. A., & Copeland, D. E. (2006). Walking through doorways causes forgetting: Situation models and experienced space. *Memory & Cognition, 34*(5), 1150-1156.

Radvansky, G. A., & Copeland, D. E. (2010). Reading times and the detection of event shift processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 210-216.

Radvansky, G. A., & Dijkstra, K. (2007). Aging and situation model processing. *Psychonomic Bulletin & Review, 14*(6), 1027-1042.

Radvansky, G. A., & Zacks, J. M. (2011). Event perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(6), 608-620.

Radvansky, G. A., & Zacks, J. M. (2014). Event cognition. Oxford University Press.

Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences, 17*, 133-140.

Radvansky, G. A., & Zacks, R. T. (1991). Mental models and the fan effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(5), 940-953.

Radvansky, G. A., Copeland, D. E., Berish, D. E., & Dijkstra, K. (2003). Aging and situation model updating. *Aging, Neuropsychology, and Cognition, 10*(2), 158-166.

Radvansky, G. A., Krawietz, S. A., & Tamplin, A. K. (2011). Walking through doorways causes forgetting: Further explorations. *The Quarterly Journal of Experimental Psychology*, *64*(8), 1632-1645.

Radvansky, G. A., Tamplin, A. K., & Krawietz, S. A. (2010). Walking through doorways causes forgetting: Environmental integration. *Psychonomic Bulletin & Review, 17*(6), 900-904.

Radvansky, G. A., Zwaan, R. A., Federico, T., & Franklin, N. (1998). Retrieval from temporally organized situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1224-1237.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience, 13*(10), 713-726.

Rauthmann, J. F. (2012). You say the party is dull, I say it is lively: A componential approach to how situations are perceived to disentangle perceiver, situation, and perceiver × situation variance. *Social Psychological and Personality Science, 3*(5), 519-528.

Rauthmann, J. F. (2015). Structuring Situational Information. European Psychologist.

Rauthmann, J. F. (2016). Motivational factors in the perception of psychological situation characteristics. *Social and Personality Psychology Compass, 10*(2), 92-108.

Rauthmann, J. F., & Sherman, R. A. (2016). Situation change: Stability and change of situation variables between and within persons. *Frontiers in Psychology, 6*, 1938.

Rauthmann, J. F., & Sherman, R. A. (2017). Normative and distinctive accuracy in situation perceptions: Magnitude and personality correlates. *Social Psychological and Personality Science, 8*(7), 768-779.

Rauthmann, J. F., & Sherman, R. A. (2018). The description of situations: Towards replicable domains of psychological situation characteristics. *Journal of Personality and Social Psychology, 114*(3), 482-488.

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., ... & Funder, D. C. (2014). The Situational Eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology, 107*(4), 677-718.

Rauthmann, J. F., Sherman, R. A., & Funder, D. C. (2015). Principles of situation research: Towards a better understanding of psychological situations. *European Journal of Personality, 29*(3), 363-381.

Reynolds, J. R., Zacks, J. M., & Braver, T. S. (2007). A computational model of event segmentation from perceptual prediction. *Cognitive Science, 31*(4), 613-643.

Rhodes, S., Cowan, N., Parra, M. A., & Logie, R. H. (2018). Interaction effects on common measures of sensitivity: Choice of measure, type I error, and power. *Behavior Research Methods,* 1-19.

Richmond, L. L., & Zacks, J. M. (2017). Constructing experience: Event models from perception to action. *Trends in Cognitive Sciences, 21*(12), 962-980.

Rinck, M., & Bower, G. H. (2000). Temporal and spatial distance in situation models. *Memory & Cognition, 28*(8), 1310-1320.

Rinck, M., & Weber, U. (2003). Who when where: An experimental test of the event indexing model. *Memory & Cognition, 31*(8), 1284-1292.

Rinck, M., Hähnel, A., Bower, G. H., & Glowalla, U. (1997). The metrics of spatial situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(3), 622-637.

Rosch, E. (1999). Principles of categorization. Concepts: core readings, 189.

Rosen, V. M., Caplan, L., Sheesley, L., Rodriguez, R., & Grafman, J. (2003). An examination of daily activities and their scripts across the adult lifespan. *Behavior Research Methods, Instruments, & Computers, 35*(1), 32-48.

Rotello, C. M., Masson, M. E., & Verde, M. F. (2008). Type I error rates and power analyses for single-point sensitivity measures. *Perception & Psychophysics, 70*(2), 389-401.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1*(1), 45-57.

Saucier, G., Bel-Bahar, T., & Fernandez, C. (2007). What modifies the expression of personality tendencies? Defining basic domains of situation variables. *Journal of Personality, 75*(3), 479-504.

Schank, R., & Abelson, R (1977). Scripts, plans, goals, and understanding. Hillsdale, N.J.: Erlbaum.

Schmalhofer, F., McDaniel, M. A., & Keefe, D. (2002). A unified model for predictive and bridging inferences. *Discourse Processes, 33*(2), 105-132.

Sherman, R. A., Nave, C. S., & Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *Journal of Personality and Social Psychology, 99*(2), 330-343.

Sherman, R. A., Nave, C. S., & Funder, D. C. (2013). Situational construal is related to personality and gender. *Journal of Research in Personality, 47*(1), 1-14.

Sherman, R. A., Rauthmann, J. F., Brown, N. A., Serfass, D. G., & Jones, A. B. (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology, 109*(5), 872-888.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors*, *37*(1), 137-148.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34-50.

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language*, *53*(1), 125-140.

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, *18*(5), 449-455.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149.

Sulistyawati, K., Wickens, C. D., & Chui, Y. P. (2011). Prediction in situation awareness: Confidence bias and underlying cognitive abilities. *The International Journal of Aviation Psychology*, *21*(2), 153-174.

Swallow, K. M., Kemp, J. T., & Simsek, A. C. (2018). The role of perspective in event segmentation. *Cognition*, *177*, 249-262.

Ten Berge, M. A., & De Raad, B. (1999). Taxonomies of situations from a trait psychological perspective. A review. *European Journal of Personality*, *13*(5), 337-360.

Tenney, Y. J., & Pew, R. W. (2006). Situation awareness catches on: what? so what? now what?. *Reviews of Human Factors and Ergonomics*, *2*(1), 1-34.

van den Broek, P., Beker, K., & Oudega, M. (2015). Inference generation in text comprehension: Automatic and strategic processes in the construction of a mental representation. Inferences during reading, 94-121.

Van Dijk, T. A., Kintsch, W., & Van Dijk, T. A. (1983). Strategies of discourse comprehension (pp. 11-12). New York: Academic Press.

Van Heck, G. L. (1984). The construction of a general taxonomy of situations. *Personality psychology in Europe: Theoretical and empirical developments*, 1, 149-164.

Van Heck, G. L. (1989). Situation concepts: Definitions and classification.

van Winsen, R., Henriqson, E., Schuler, B., & Dekker, S. W. (2015). Situation awareness: some conditions of possibility. *Theoretical Issues in Ergonomics Science, 16*(1), 53-68.

Verde, M. F., Macmillan, N. A., & Rotello, C. M. (2006). Measures of sensitivity based on a single hit rate and false alarm rate: The accuracy, precision, and robustness of d′, A_z, and A'. *Perception & Psychophysics, 68*(4), 643-654.

Wassenburg, S. I., Bos, L. T., de Koning, B. B., & van der Schoot, M. (2015). Effects of an inconsistency-detection training aimed at improving comprehension monitoring in primary school children. *Discourse Processes, 52*(5-6), 463-488.

Wickens, C. D. (2015). Situation awareness: Its applications value and its fuzzy dichotomies. *Journal of Cognitive Engineering and Decision Making, 9*(1), 90-94.

Wilder, D. A. (1978a). Effect or predictability on units of perception and attribution. *Personality and Social Psychology Bulletin, 4*, 281–284

Wilder, D. A. (1978b). Predictability of behaviors, goals, and unit of perception. *Personality and Social Psychology Bulletin, 4*(4), 604-607.

Woods, D. D. (1988, March). Coping with complexity: The psychology of human behaviour in complex systems. In Tasks, errors, and mental models (pp. 128-148). Taylor & Francis, Inc.

Wyer Jr, R. S., & Radvansky, G. A. (1999). The comprehension and validation of social information. *Psychological Review, 106*(1), 89-118.

Yang, Y., Read, S. J., & Miller, L. C. (2006). A taxonomy of situations from Chinese idioms. *Journal of Research in Personality, 40*(5), 750-778.

Yang, Y., Read, S. J., & Miller, L. C. (2009). The concept of situations. *Social and Personality Psychology Compass, 3*(6), 1018-1037.

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28*(6), 979-1008.

Zacks, J. M., & Magliano, J. P. (2011). Film, narrative, and cognitive neuroscience. Art and the senses, 435-454.

Zacks, J. M., & Sargent, J. Q. (2010). Event perception: A theory and its application to clinical neuroscience. In Psychology of learning and motivation (Vol. 53, pp. 253-299). Academic Press.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3-21.

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., ... & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*(6), 651-655.

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition, 112*(2), 201-216.

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *Journal of Cognitive Neuroscience*, *23*(12), 4057-4066.

Zacks, J. M., Kurby, C. A., Landazabal, C. S., Krueger, F., & Grafman, J. (2016). Effects of penetrating traumatic brain injury on event segmentation and memory. *Cortex, 74*, 233-246.

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General, 138*(2), 307-327.

Zacks, J. M., Speer, N. K., Swallow, K. M., & Maley, C. J. (2010). The brain's cutting-room floor: Segmentation of narrative cinema. *Frontiers in Human Neuroscience, 4*, 168.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin, 133*(2), 273-293.

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging, 21*(3), 466-482.

Zacks, J. M., Swallow, K. M., Vettel, J. M., & McAvoy, M. P. (2006). Visual motion and the neural correlates of event perception. *Brain Research, 1076*(1), 150-162.

Zalla, T., Labruyère, N., & Georgieff, N. (2013). Perceiving goals and actions in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*(10), 2353-2365.

Zalla, T., Pradat-Diehl, P., & Sirigu, A. (2003). Perception of action boundaries in patients with frontal lobe damage. *Neuropsychologia, 41*(12), 1619-1627.

Zalla, T., Verlut, I., Franck, N., Puzenat, D., & Sirigu, A. (2004). Perception of dynamic action in patients with schizophrenia. *Psychiatry Research, 128*(1), 39-51.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1196-1207.

Zwaan, R. A. (1999). Situation models: The mental leap into imagined worlds. *Current Directions in Psychological Science, 8*(1), 15-18.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162-185.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*(5), 292-297.

Zwaan, R. A., Madden, C. J., & Whitten, S. N. (2000). The presence of an event in the narrated situation affects its availability to the comprehender. *Memory & Cognition*, *28*(6), 1022-1028.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 386-397.

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading, 2*(3), 199-220.